# [0001] RECOMBINATION SYSTEMS AND METHODS FOR ELIMINATING NUCLEIC ACID SEQUENCES FROM THE GENOME OF EUKARYOTIC ORGANISMS

[0002] **Background**

[0003] **1. Field of the Invention**

[0004] The invention relates to recombination systems and methods for eliminating nucleic acid sequences from the genome of eukaryotic organisms, and to transgenic organisms – preferably plants – comprising these systems.

[0005] **2. Description of the Background**

[0006] The purpose of biotechnological research into organisms consists in, inter alia, obtaining commercially utilizable information on the function of certain genes and gene products and the elucidation of biosynthetic pathways or disease mechanisms. The information obtained in this manner can be employed in a multiplicity of ways. They serve for example for the production of novel medicaments, the development of alternative, biotechnological production methods or the generation of modified plants. An aim of biotechnological research into plants is the generation of plants with advantageous novel characteristics, for example for increasing agricultural productivity, improving the quality in foodstuffs or for the production of certain chemicals or pharmaceuticals (Dunwell JM, J Exp Bot. 2000;51 Spec No:487-96).

[0007] In the generation of transgenic organisms, selection of the organisms which have been modified in the desired manner is required owing to the poor efficacy of the methods used (such as, for example, stable transformation or, in particular, homologous recombination). Transgenic plants can be generated by a series of techniques (Review: Potrykus I. and Spangenberg G. ed. (1995) *Gene transfer to plants*. Springer, Berlin). In particular the gene transfer mediated by *Agrobacterium tumefaciens* and the bombardment of plant cells with the particle gun play an important role in this context. An important problem is the fact that transgenic DNA, once stably introduced into an organism, can only be removed with difficulty. The genes for resistance to antibiotics or herbicides, which are used during the transformation procedure for selection purposes, remain in the transgenic

plants, which contributes substantially to the lack of acceptance of these "gene food" products among consumers.

[0008]    It has therefore been attempted for some time to develop techniques by means of which foreign DNA can be integrated into the plant genome at the specific sites or reexcised therefrom (Ow DW and Medberry SL (1995) Crit Rev in Plant Sci 14:239-261).

[0009]    The skilled worker is familiar with a variety of systems for the site-directed removal of recombinantly introduced nucleic acid sequences. They are based on the use of sequence—specific recombinases and two recognition sequences of said recombinases which flank the sequence to be removed. The effect of the recombinase on this construct brings about the excision of the flanked sequence, one of the recognition sequences remaining in the genome of the organism. Various sequence-specific recombination systems are described, such as the Cre/lox system of the bacteriophage P1 (Dale EC and Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562; Russell SH et al. (1992) Mol Gene Genet 234: 49-59; Osborne BI et al. (1995) Plant J. 7, 687-701), the yeast FLP/FRT system (Kilby NJ et al. (1995) Plant J 8:637-652; Lyznik LA et al. (1996) Nucleic Acids Res 24:3784-3789), the Mu phage Gin recombinase, the E. coli Pin recombinase or the R/RS system of the plasmid pSR1 (Onouchi H et al.(1995) Mol. Gen. Genet. 247:653-660; Sugita K et al. (2000) Plant J. 22:461-469). Here, the recombinase (for example Cre or FLP) interacts specifically with its corresponding recombination sequences (34 bp lox sequence and 47 bp FRT sequence, respectively) in order to delete or invert the interposed sequences. Reports on successful applications of these systems in plants are limited. Thus, David Ow's group has demonstrated that a selection marker used for the transformation of plants which was flanked by two lox sequences can be reexcised from the plant genome by the expression of Cre (Dale EC and Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562). A disadvantage of the sequence-specific recombination systems is the reversibility of the reaction, that is to say an equilibrium exists between excision and integration of the marker gene in question. This frequently brings about the selection of mutations, i.e. as soon as a mutation blocks the further interaction of the lox recognition sequences with the enzyme,

va-53897

the (undesired) product is removed from the equilibrium and fixed. This not only applies to the Cre-lox system, but also to the other sequence-specific recombinases (see above). A further disadvantage is the fact that one of the recognition sequences of the recombinase remains in the genome, which is thus modified. This may have effects on the characteristics of the organisms when, for example, the recognition sequence modifies or destroys reading frames or genetic control elements such as promotors or enhancers. Furthermore, the recognition sequence which remains in the genome excludes a further use of the recombination system, for example for a second genetic modification, since interactions with the subsequently introduced recognition sequences cannot be ruled out. Substantial chromosomal rearrangements or deletions may result.

[0010]   Zubko et al. describe a system for the deletion of nucleic acid sequences from the tobacco genome, where the sequence to be deleted is flanked by two 352 bp attP recognition sequences from the bacteriophage Lambda. Deletion of the flanked region takes place independently of the expression of helper proteins in two of eleven transgenic tobacco lines by spontaneous intrachromosomal recombination between the attP recognition regions. The disadvantages of this method are that recombination, or deletion, cannot be induced specifically at a particular point in time, but takes place spontaneously. The fact that the method worked only in a small number of lines suggests that the integration locus in the cases in question tends to be unstable (Puchta H (2000) Trends in Plant Sci 5:273-274).

[0011]   On page 12 in the key to Figure 32, WO 96/14408 describes a method for eliminating a genetic locus in which in each case one recognition sequence of the homing restriction endonuclease I-SceI is inserted at the respective end of the sequence to be deleted. Treatment with the endonuclease leads to double-strand breaks at both ends of the sequence to be deleted. The free ends then join up by means of "recombination". The "recombination" cited here can only be an illegitimate recombination – as can also be seen from the Figure - (for example a non-homologous end-joining (NHEJ) event), since no homologous sequences exist at the two remaining ends of the genomic DNA. Illegitimate

va-53897

3

recombination, however, leads to unpredictable recombination events. This may have effects on the characteristics of the organisms if for example reading frames or genetic control elements such as promotors or enhancers are modified or destroyed thereby. The system requires two recognition sequences which flank the fragment to be deleted.

[0012] The generation of sequence-specific double-strand breaks with the aid of restriction enzymes in eukaryotic genomes such as yeast (Haber JE (1995) Bioassays 17:609-620), mammalian cells (Jasin M (1996) Trends Genet. 12:224-228) or plants (Puchta H (1999a) Methods Mol Biol 113:447-451) is described.

[0013] What is described is the induction of an intramolecular recombination on a plasmid DNA in Xenopus oocytes by sequence-specific cleavage with the endonuclease I-SceI (Segal DJ and Caroll D (1995) Proc Natl Acad Sci USA 92:806-810) or by synthetic, chimeric nucleases (Bibikova M et al. (2001) Mol Cell Biol 21(1):289-297). The aim is the site-directed recombination between two homologous sequences between which a suitable nuclease cleavage site is located. Both cases are extrachromosomal recombination events in which in each case only part of the extra chromosomal plasmid DNA undergoes homologous recombination.

[0014] Posfai et al. describe a method for exchanging genes in the prokaryote E.coli (Posfai G et al. (1999) Nucleic Acids Res 27(22):4409-4415). Here, recombination between the endogenous and the mutated gene results in the E.coli genome, induced by cleavage with the restriction enzyme I-SceI. Aim and object was the exchange of an endogenous gene for a mutated transgene. Recombinations in E.coli proceed in a markedly simpler way and with greater efficacy than in higher eukaryotes (for example described by Kuzminov A (1999) Microbiol Mol Biol Rev. 63(4):751-813).

[0015] Dürrenberger et al. describe the induction of recombination in chloroplasts of the single-celled green alga Chlamydomonas reinhardtii using the I-SceI homing endonuclease (Dürrenberger F et al. (1996) Nucleic Acid Res 24(17):3323-3331). Recombination takes place between the endogenous 23S gene and an inserted 23S cDNA which contains a I-SceI cleavage site. Double-strand breaks are induced by mating the transgenic organism in

va-53897

question with an organism expressing I-SceI. Recombinations in chloroplasts proceed in a markedly simpler manner and with greater efficacy than in the chromosomal DNA of higher eukaryotes. Thus, indeed, homologous recombination appears to be the preferred, normal way of DNA integration in plastids (chloroplasts) (described in: Heifetz PB and Tuttle AM (2001) Curr Opinion Plant Biol 4:157-161). It appears that plastids have a specific system which enables them to undergo homologous recombination, as opposed to the nucleus, and facilitates the site-directed introduction of foreign DNA (Heifetz PB (2000) Biochimie 82:655-666).

[0016]     The gene targeting technique, in which a site-directed integration into the chromosomal DNA of the host organism is to be achieved by means of homologous recombination works acceptably well only in the case of prokaryotes and yeast. The generation of corresponding transgenic organisms is possible in a few species only (such as, for example, mice) and even then highly complicated (see also Kanaar R Hoeijmakers JH (1997) Genes Funct 1(3):165-174). The existing, poor homologous recombination efficacy (approx. $1:1x10^6$) is compensated for in this case by the use of complicated, sophisticated selection techniques which are limited to the species in question (such as, for example, "ES" cell technology). In other species – but above all in Higher Plants – such technologies have not been established as yet (Mengiste T and Paszkowski J (1999) Biol Chem. 380:749-758; Vergunst AC and Hooykaas PJJ (1999) Crit Rev Plant Sci 18:1-31; Puchta H (1999) Methods Mol Biol 113:447-451; Hohn B and Puchta H (1999) Proc Natl Acad Sci USA 96:8321-8323). Attempts to achieve homologous recombination in plants resulted in random, nonhomologous "illegitimate" insertion events in most cases. Here, the DNA introduced is integrated at one or more unpredictable sites in the plant genome. Integration takes place by means of illegitimate recombination (Roth DB and Wilson JH (1988) illegitimate recombination in mammalian cells. In "Genetic recombination", R. Kucherlapati and G.R. Smith Edts., American Society of Microbiology, Washington, USA; pp.621-635) and not in sequence regions which are homologous to the transferred DNA (Puchta H and Hohn B (1996) Trends Plant Sci. 1:340-348). The problem of lacking

va-53897

5

efficacy in homologous recombination, which is serious predominantly in plants, is generally known to the skilled worker. The causes are addressed by current research (Review article: Mengiste T and Paszkowski J (1999) Biological Chemistry 380(7-8):749-58). Increasing the efficacy of homologous recombination has long been a need in plant biotechnology which is hitherto unresolved.

[0017]     A further need which has long existed in biotechnological research and which is not addressed by any of the established systems is the provision of systems and methods which enables the site-directed elimination of nucleic acid sequences from the chromosomal DNA of a eukaryotic organism and allow the repeated application to the same organism. For example, it is an aim of plant biotechnology further to improve by means of recombinant methods existing high-yielding varieties. In this context, it is particularly important to eliminate, after the transformation has taken place, superfluous transgene sequences such as selection markers. In addition, methods for the predictable elimination of sequences, for example from the chromosomal DNA of an organism, would offer further applications in the field of genetic engineering which are of great interest scientifically and economically.

[0018]     **Summary of the Invention**

[0019]     The present invention is directed to systems and methods that enable the predictable elimination of defined nucleic acid sequences from the chromosomal DNA of a eukaryotic organism and allow the repeated, successive application to the same organism. This invention was achieved in a surprising manner by providing the recombination system according to the invention.

[0020]     One embodiment of the invention is directed to recombination systems for eliminating DNA sequences from the chromosomal DNA of eukaryotic cells or organisms, which comprises, in a eukaryotic cell or organism, a transgenic recombination construct inserted into the chromosomal DNA of a eukaryotic organism comprising a sequence · consisting, in the 5'/3'-direction, of a first homology sequence A and at least one recognition sequence for the site-directed induction of DNA double-strand breaks and a

va-53897

second homology sequence B, the homology sequences A and B having a sufficient length and sufficient homology in order to ensure homologous recombination, together with an enzyme suitable for inducing DNA double-strand breaks at the recognition sequence (b1) for the site-directed induction of DNA double-strand breaks or a nucleic acid sequence encoding an enzyme suitable for inducing DNA double-strand breaks at the recognition sequence (b1).

[0021]    Another embodiment of the invention is directed to methods for eliminating DNA sequences from chromosomal DNA of eukaryotic cells or organisms, which comprises combining, in a eukaryotic cell or organism, a transgenic recombination construct inserted into the chromosomal DNA of a eukaryotic organism comprising a sequence consisting, in the 5'- to 3'-orientation, of a first homology sequence A and at least one recognition sequence for the site-directed induction of DNA double-strand breaks and a second homology sequence B, the homology sequences A and B having a sufficient length and sufficient homology in order to ensure homologous recombination, together with an enzyme suitable for inducing DNA double-strand breaks at the recognition sequence (b1) for the site-directed induction of DNA double-strand breaks, and the induction of DNA double-strand breaks at the recognition sequence for the site-directed induction of DNA double-strand breaks and the homologous recombination taking place between the homology sequences A and B.

[0022]    Other embodiments and advantages of the invention are set forth in part in the description, which follows, and in part, may be obvious from this description, or may be learned from the practice of the invention.

[0023]    **Description of the Figures**

[0024]    Figure 1. Diagram of a principal of the invention.

[0025]    Figure 2. Diagram of a preferred embodiment of the invention.

[0026]    Figure 3. Diagram of another preferred embodiment of the invention.

[0027]    Figure 4. Diagram of another preferred embodiment of the invention.

[0028]    Figure 5. Diagram of another preferred embodiment of the invention.

va-53897

[0029] Figure 6. Diagram of another preferred embodiment of the invention.

[0030] Figure 7. Diagrams of exemplary embodiments of the invention.

[0031] Figure 8. Diagram of another preferred embodiment of the invention.

[0032] Figure 9. Readily selectable systems for deleting nucleic acid sequences from chromosomal DNA of an organism.

[0033] Figure 10. Two constructs that demonstrate homologous recombination by double-strand breaks can be induced with restriction enzymes.

[0034] Figure 11. Representative histochemical analysis of tobacco calli obtained after induction of double-strand breaks.

[0035] Figure 12. PCR analysis for detecting homologous recombination.

[0036] Figure 13. Southern blots indicating complete elimination of a transgene sequence.

[0037] **Description of the Invention**

[0038] As embodied and broadly described herein, the present invention is directed to systems and methods that enable the predictable elimination of defined nucleic acid sequences from the chromosomal DNA of a eukaryotic organism. The invention enables sequences (for example selection markers such as genes for resistance to antibiotics or herbicides) to be deleted from the chromosomal DNA of an organism in an accurately predictable manner. In doing so, the sequence to be eliminated is flanked by recognition sequences for the site-directed induction of DNA double-strand breaks (for example recognition sequences of rare-cleaving restriction enzymes) and combined with homologous sequences in the region of the cleavage sites. A double-strand break is induced by an enzyme suitable for inducing DNA double-strand breaks at the recognition sequence for the site-directed induction of DNA double-strand breaks (for example a sequence-specific nuclease), which, in consequence, triggers the homologous recombination of homologous sequences located at the break, and thus the deletion of any nucleic acid sequences located between the sequences. The recognition sequence for the site-directed

induction of DNA double-strand breaks is likewise deleted, and the method can thus be used repeatedly for further controlled genetic modifications.

[0039] Surprisingly, this induced homologous recombination takes place with high efficacy and precision, which is in contrast to previous experience in the field of homologous recombination, including in plants. The frequency can be compared with the parallel, nonhomologous events (for example non-homologous end-joining events) (cf. Example 5). This is a remarkable finding which is in contrast to earlier observations, according to which the frequency of homologous recombination – above all in the case of plants – is secondary, almost negligible, in comparison with the "illegitimate" events.

[0040] The sequences which are deleted are those located between the homology sequences A and B. In contrast to systems such as, for example, the cre/lox or the FRT/FLP system, one is not bound to specific sequences when performing recombination. The skilled worker knows that any sequence can undergo homologous recombination with another sequence provided that sufficient length and homology exist. Owing to the sequence-specific induction of the double-strand breaks, the homologous recombination efficacy between the homology sequences A and B is increased considerably, indeed enabled in the first place in some cases.

[0041] With regard to the recombination construct, "transgene" refers to all those constructs which are the result of recombinant methods in which either

[0042] a) at least one of the homology sequences A or B, or

[0043] b) at least one recognition sequence for the site directed induction of DNA double-strand breaks, or

[0044] c) (a) and (b)

[0045] are not located in their natural genetic environment (for example at their natural chromosomal locus) or have been modified by recombinant methods, it being possible for the modification to encompass, for example, substitutions, additions, deletions, inversion or insertions of one or more nucleotide residues.

va-53897

9

[0046] "Eukaryotic cell or organism" generally refers to any eukaryotic cell or organism and to cells, tissues, parts or propagation material (such as seeds or fruits) derived from these in which an induction of double-strand breaks may take place at the recognition sequence for the site-directed induction of DNA double-strand breaks and the homologous recombination between the homology sequences A and B may take place while the recombination construct and the enzyme suitable for inducing DNA double-strand breaks at the recognition sequence for the site-directed induction of DNA double-strand breaks are simultaneously present in one reaction space (for example in a cell or cell compartment). A particularly preferred embodiment encompasses compartments of a eukaryotic cell such as, for example, the nucleus.

[0047] Cells or organisms that are especially preferably encompassed are those which constitute a multi-celled eukaryotic organism or are derived from the latter, and cells, tissues, parts or propagation material (such as seeds or fruits) of same. Very especially preferably encompassed cells or organisms are those which constitute an animal or plant organism or are derived from the latter, and cells, tissues, parts or propagation material of same. Most preferably encompassed cells or organisms are those which constitute a plant organism or are derived from the latter, and cells, tissues, parts or propagation material of same. Preferred genera and species are detailed further below.

[0048] Referring to the homology sequences A and B, "sufficient length" preferably refers to sequences with a length of at least 20 base pairs, preferably at least 50 base pairs, especially preferably at least 100 base pairs, very especially preferably at least 250 base pairs, most preferably at least 500 base pairs.

[0049] Referring to the homology sequences A and B, "sufficient homology" preferably refers to sequences with at least 70%, preferably 80%, by preference at least 90%, especially preferably at least 95%, very especially preferably at least 99%, most preferably 100%, homology within these homology sequences over a length of at least 20 base pairs, preferably at least 50 base pairs, especially preferably at least 100 base pairs, very especially preferably at least 250 base pairs, most preferably at least 500 base pairs.

va-53897

10

[0050]     Homology between two nucleic acid sequences is understood as meaning the identity of the nucleic acid sequence over in each case the entire sequence length which is calculated by alignment with the aid of the program algorithm GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

[0051]          Gap Weight: 12          Length Weight: 4

[0052]          Average Match: 2,912          Average Mismatch:-2,003

[0053]     In a preferred embodiment, only one recognition sequence for the site-directed induction of DNA double-strand breaks is located between the homology sequences A and B, so that the recombination construct employed in the recombination system or method according to the invention is constructed in the 5'- to 3'-orientation as follows:

[0054]     a1) a first homology sequence A, and

[0055]     b1) a recognition sequence for the site-directed induction of DNA double-strand breaks, and

[0056]     a2) a second homology sequence B, the homology sequences A and B having a sufficient length and sufficient homology in order to ensure homologous recombination.

[0057]     In a preferred embodiment, a further nucleic acid sequence is located between the homology sequences A and B, so that the recombination construct employed in the recombination system or method according to the invention is constructed as follows in the 5'/3'-direction of:

[0058]     a1) a first homology sequence A, and

[0059]     b1) a recognition sequence for the site-directed induction of DNA double-strand breaks, and

[0060]     c)  a further nucleic acid sequence, and

[0061]     a2) a second homology sequence B, the homology sequences A and B having a sufficient length and sufficient homology in order to ensure homologous recombination.

[0062]     The recognition sequence for the site-directed induction of DNA double-strand breaks may also be located after or within the further nucleic acid sequence.

va-53897

11

[0063]    In a further preferred embodiment, a second recognition sequence for the site-directed induction of double-strand breaks is present after the further nucleic acid sequence. This embodiment is advantageous in particular in the case of homology sequences A and B which are further apart, or in the case of longer further nucleic acid sequences, since recombination efficacy is increased. In this embodiment, the recombination construct employed in the recombination system or method according to the invention is constructed as follows in a 5'- to 3'-orientation of:

[0064]    a1) a first homology sequence A, and

[0065]    b1) a first recognition sequence for the site-directed induction of DNA double-strand breaks, and

[0066]    c) a further nucleic acid sequence, and

[0067]    b2) a second recognition sequence for the site-directed induction of DNA double-strand breaks, and

[0068]    a2) a second homology sequence B, the homology sequences A and B having a sufficient length and sufficient homology in order to ensure homologous recombination.

[0069]    Furthermore, other recognition sequences may also be present between the homology sequences A and B, in addition to the second recognition sequences for the site-directed induction of DNA double-strand breaks. The individual recognition sequences (for example b1 or b2) for the site-directed induction of DNA double-strand breaks may be identical or different, i.e. they may act as recognition sequence for an individual enzyme for the site-directed induction of DNA double-strand breaks or else for a variety of enzymes. Here, the embodiment in which the recognition sequences for the site-directed induction of DNA double-strand breaks act as recognition sequence for an individual enzyme for the site-directed induction of DNA double-strand breaks is preferred in this context.

[0070]    The skilled worker is familiar with a variety of ways to obtain one of the recombination constructs according to the invention. They can be prepared by means of customary recombination and cloning techniques as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold

va-53897

12

Spring Harbor Laboratory, Cold Spring Harbor, NY (1989), in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987).

[0071] Preferably, the recombination construct according to the invention is generated by joining the abovementioned essential constituents of the recombination construct together in the abovementioned sequence using the recombination and cloning techniques with which the skilled worker is familiar, and the result is then introduced into the chromosomal DNA of a host organism.

[0072] However, the skilled worker is aware that he may also obtain the recombination construct according to the invention in other ways. Thus, the host organism may already comprise one or more of the essential components of the recombination construct. The recombination construct according to the invention is then generated by introducing one further, or more, essential components of the recombination construct in the correct position relative to the existing components in said organism. Thus, for example, the starting organism may already comprise one of the homology sequences A or B. If the organism already comprises a homology sequence A, introducing a construct consisting of a recognition sequence for the site-directed induction of DNA double-strand breaks and a second homology sequence B after the homology sequence A gives rise to one of the recombination constructs according to the invention.

[0073] Furthermore, the skilled worker is familiar with various ways in which the recombination construct according to the invention may be introduced into the chromosomal DNA of a eukaryotic cell or organism. In this context, the insertion may be directed (i.e. taking place at a defined insertion site) or undirected (i.e. taking place randomly). Suitable techniques are known to the skilled worker and described by way of example herein.

[0074] "Enzyme suitable for inducing DNA double-strand breaks in the recognition sequence for the site-directed induction of DNA double-strand breaks" (herein "DSBI

va-53897

13

enzyme", which stands for "double strand-break inducing enzyme") generally refers to all those enzymes which are capable of generating double-strand breaks in double stranded DNA in a sequence-specific manner. The following may be mentioned by way of example, but not by limitation:

[0075]    1.  Restriction endonucleases (type II), preferably homing endonucleases as described in detail herein.

[0076]    2.  Recombinases (such as, for example, Cre/lox; R-RS; FLP/FTR as described above).

[0077]    3.  Transposases, for example the P-element transposase (Kaufman PD and Rio DC (1992) Cell 69(1):27-39) or AcDs (Xiao YL and Peterson T (2000) Mol Gen Genet 263(1):22-29). In principle, all transposases or integrases are suitable as long as they have sequence specificity (Haren L et al. (1999) Annu Rev Microbiol. 1999;53:245-281; Beall EL, Rio DC (1997) Genes Dev. 11(16):2137-2151).

[0078]    4.  Chimeric nucleases as described in detail herein.

[0079]    5.  Enzymes which induce double-strand breaks in the immune system, such as the RAG1/RAG2 system (Agrawal A et al. (1998) Nature 394(6695):744-451).

[0080]    6.  Group II intron endonucleases. Modifications of the intron sequence allows group II introns to be directed to virtually any sequence in a double-stranded DNA, where group II introns can subsequently insert by means of a reverse splice mechanism (Mohr et al. (2000) Genes & Development 14:559-573; Guo et al. (2000) Science 289:452- 457). During this reverse splice mechanism, a double-strand break is introduced into the target DNA, the excised intron RNA cleaving the sense strand while the protein portion of the group II intron endonuclease hydrolyses the antisense strand (Guo et al. (1997) EMBO J 16: 6835- 6848). If it is only desired to induce the double-strand break without achieving complete reverse splicing, as is the case in the present invention, it is possible to resort to, for example, group II intron endonucleases which lack the reverse transcriptase activity. While this does not prevent the generation of the double-strand break, the reverse splicing

mechanism cannot proceed to completion. Suitable enzymes are not only natural enzymes, but also synthetic enzymes.

[0081]     Preferred enzymes are all those DSBI enzymes whose recognition sequence is known and which can either be obtained in the form of their proteins (for example by purification) or expressed using their nucleic acid sequence.  Especially preferred are restriction endonucleases (restriction enzymes) which have no or only a few recognition sequences – besides the recognition sequences present in the transgenic recombination construct - in the chromosomal DNA sequence of a particular eukaryotic organism. This avoids further double-strand breaks at undesired loci in the genome.

[0082]     This is why homing endonucleases are very especially preferred (Review: (Belfort M and Roberts RJ (1997) Nucleic Acids Res 25: 3379-3388; Jasin M (1996) Trends Genet. 12:224-228; Internet: http://rebase.neb.com/rebase/rebase.homing.html). Owing to their long recognition sequences, they have no, or only a few, further recognition sequences in the chromosomal DNA of eukaryotic organisms in most cases.

[0083]     The sequences encoding for such homing endonucleases can be isolated for example from the chloroplast genome of Chlamydomonas (Turmel M et al. (1993) J Mol Biol 232: 446-467). They are small (18 to 26 kD) and their open reading frame (ORF) has a "coding usage" which is suitable directly for nuclear expression in eukaryotes (Monnat RJ Jr et al. (1999) Biochem Biophys Res Com 255:88-93). Homing endonucleases which are very especially preferably isolated are the homing endonucleases I-SceI (WO96/14408), I-SceII (Sarguiel B et al. (1990) Nucleic Acids Res 18:5659-5665), I-SceIII (Sarguiel B et al. (1991) Mol Gen Genet. 255:340-341), I-CeuI (Marshall (1991) Gene 104:241-245), I-CreI (Wang J et al. (1997) Nucleic Acids Res 25: 3767-3776), I-ChuI (Cote V et al.(1993) Gene 129:69-76), I-TevI (Chu et al. (1990) Proc Natl Acad Sci USA 87:3574-3578; Bell-Pedersen et al. (1990) Nucleic Acids Res18:3763-3770), I-TevII (Bell-Pedersen et al. (1990) Nucleic Acids Res18:3763-3770), I-TevIII (Eddy et al. (1991) Genes Dev. 5:1032-1041), Endo SceI (Kawasaki et al. (1991) J Biol Chem 266:5342-5347), I-CpaI (Turmel M

et al. (1995a) Nucleic Acids Res 23:2519-2525) and I-CpaII (Turmel M et al. (1995b) Mol. Biol. Evol. 12, 533-545).

[0084] Further homing endonucleases are detailed in the abovementioned Internet website, and examples which may be mentioned are homing endonucleases such as F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-AmaI, I-AniI, I-CeuI, I-CeuAIIP, I-ChuI, I-CmoeI, I-CpaI, I-CpaII, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdiI, I-DdiII, I-DirI, I-DmoI, I-HmuI, I-HmuII, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII , I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiS3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPA1P, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-PspI, PI-Rma43812IP, PI-SPBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII, and combinations thereof.

[0085] Preferred in this context are the homing endonucleases whose gene sequences are already known, such as, for example, F-SceI,I-CeuI,I-ChuI, I-DmoI, I-CpaI, I-CpaII, I-CreI, I-CsmI, F-TevI, F-TevII, I-TevI, I-TevII, I-AniI, I-CvuI, I-DdiI, I-HmuI, I-HmuII, I-LlaI, I-NanI, I-MsoI, I-NitI, I-NjaI, I-PakI, I-PorI, I-PpoI, I-ScaI, I-Ssp6803I, PI-PkoI, PI-PkoII, PI-PspI, PI-TfuI, PI-TliI, and combinations thereof.

[0086] Very especially preferred are commercially available homing endonucleases such as I-CeuI, I-SceI, I-DmoI, I-PpoI, PI-PspI or PI-SceI.

[0087] The enzymes can be isolated from their organisms of origin in the manner with which the skilled worker is familiar, and/or their coding nucleic acid sequence can be cloned. The sequences of various enzymes are deposited in GenBank.

[0088] Very especially preferred are the homing endonucleases I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI. Most preferred are the homing endonucleases as shown in SEQ ID NO: 2, 4, 6, 8 or 10, or combinations thereof.

va-53897

[0089] Synthetic DSBI enzymes which may be mentioned by way of example are chimeric nucleases which are composed of an unspecific nuclease domaine and a sequence-specific DNA binding domaine consisting of zinc fingers (Bibikova M et al. (2001) Mol Cell Biol. 21:289-297). These DNA-binding zinc finger domains can be adapted to suit any DNA sequence. Suitable methods for preparing suitable zinc finger domains are described and known to the skilled worker (Beerli RR et al., Proc Natl Acad Sci U S A. 2000; 97 (4):1495-1500; Beerli RR, et al., J Biol Chem 2000; 275(42):32617-32627; Segal DJ and Barbas CF 3rd., Curr Opin Chem Biol 2000; 4(1):34-39; Kang JS and Kim JS, J Biol Chem 2000; 275(12):8742-8748; Beerli RR et al., Proc Natl Acad Sci USA 1998; 95(25):14628-14633; Kim JS et al., Proc Natl Acad Sci USA 1997; 94(8):3616-3620; Klug A, J Mol Biol 1999; 293(2):215-218; Tsai SY et al., Adv Drug Deliv Rev 1998;30(1-3):23-31; Mapp AK et al., Proc Natl Acad Sci USA 2000; 97(8):3930-3935; Sharrocks AD et al., Int J Biochem Cell Biol 1997; 29(12):1371-1387; Zhang L et al., J Biol Chem 2000; 275(43):33850-33860).

[0090] The DSBI enzyme is preferably expressed as a fusion protein with a nuclear localization sequence (NLS). This NLS sequence enables facilitated transport into the nucleus and increases the efficacy of the recombination system. A variety of NLS sequences are known to the skilled worker and described, inter alia, by Jicks GR and Raikhel NV (1995) Annu. Rev. Cell Biol. 11:155-188. Preferred for plant organisms is, for example, the NLS sequence of the SV40 large antigen. Very especially preferred are the following NLS sequences:

[0091] NLS1: N-Pro-Lys-Thr-Lys-Arg-Lys-Val-C (SEQ ID NO: 29)

[0092] NLS2: N-Pro-Lys-Lys-Lys-Arg-Lys-Val-C (SEQ ID NO: 30)

[0093] The homing endonucleases as shown in SEQ ID NO: 4, 6, 8 or 10 used in the use examples are fusion proteins of the native nucleases and the NLS2 nuclear localization sequence.

[0094] Owing to the small size of many DSBI enzymes (such as, for example, the homing endonucleases), an NLS sequence is not necessarily required. These enzymes are

va-53897

17

capable of passing through the nuclear pores even without any aid. This is confirmed by the efficacy of the homing endonuclease as shown in SEQ ID NO: 2 which has been used and which encompasses no nuclear localization sequence.

[0095]     In a further preferred embodiment, the activity of the DSBI enzyme can be induced. Suitable methods have been described for sequence-specific recombinases (Angrand PO et al. (1998) Nucl. Acids Res. 26(13):3263-3269; Logie C and Stewart AF (1995) Proc Natl Acad Sci USA 92(13):5940-5944; Imai T et al. (2001) Proc Natl Acad Sci USA 98(1):224-228). These methods employ fusion proteins of the DSBI enzyme and the ligand binding domain for steroid hormone receptor (for example the human androgen receptor, or mutated variants of the human estrogen receptor as described therein). Induction may be effected with ligands such as, for example, estradiol, dexamethasone, 4-hydroxytamoxifen or raloxifen.

[0096]     Some DBSI enyzmes are active as dimers (homo- or heterodimers; I-CreI forms a homodimer; I-SecIV forms a heterodimer) (Wernette CM (1998) Biochemical & Biophysical Research Communications 248(1):127-333)). Dimerization can be designed as an inducible feature, for example by exchanging the natural dimerization domains for the binding domaine of a low-molecular-weight ligand. Addition of a dimeric ligand then brings about dimerization of the fusion protein. Corresponding inducible dimerization methods, and the preparation of the dimeric ligands, have been described (Amara JF et al. (1997) Proc Natl Acad Sci USA 94(20): 10618-1623; Muthuswamy SK et al. (1999) Mol Cell Biol 19(10):6845-685; Schultz LW and Clardy J (1998) Bioorg Med Chem Lett. 8(1):1-6; Keenan T et al. (1998) Bioorg Med Chem. 6(8):1309-1335).

[0097]     "Recognition sequence for the site-directed induction of DNA double-strand breaks" generally refers to those sequences which, under the conditions in the eukaryotic cell or organism used in each case, enable the recognition and cleavage by the DSBI enzyme. The recognition sequences for the respective DSBI enzymes detailed are mentioned in Table 1 herein by way of example, but not by limitation.

va-53897

[0098]     Table 1: Recognition sequences and organisms of origin of DSBI enyzmes ("^" indicates the cleavage site of the DSBI enzyme within a recognition sequence).

| DSBI enzyme | Organism of origin | Recognition sequence |
|---|---|---|
| CRE | Bacteriophage P1 | 5'-AACTCTCATCGCTTCGGATAACTTCCTGTTATCCGAAACATATCACTCACTTTGGTGATTTCACCGTAACTGTCTATGATTAATG-3' |
| FLP | Saccharomyces cerevisiae | 5'-GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC-3' |
| R | pSR1 Plasmids | 5'-CGAGATCATATCACTGTGGACGTTGATGAAAGAATACGTTATTCTTTCATCAAATCGT |
| P-Element Trans-posase | Drosophila | 5'-CTAGATGAAATAACATAAGGTGG |
| I-AniI | Aspergillus nidulans | 5'-TTGAGGAGGTT^TCTCTGTAAATAANNNNNNNNNNNNNNNN<br>3'-AACTCCTCCAAAGAGACATTTATTNNNNNNNNNNNNNNNN^ |
| I-DdiI | Dictyostelium discoideumAX3 | 5'-TTTTTTGGTCATCCAGAAGTATAT<br>3'-AAAAAACCAG^TAGGTCTTCATATA |
| I-CvuI | Chlorella vulgaris | 5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG<br>3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC |
| I-CsmI | Chlamydomonas smithii | 5'-GTACTAGCATGGGGTCAAATGTCTTTCTGG |
| I-CmoeI | Chlamydomonas moewusii | 5'-TCGTAGCAGCT^CACGGTT<br>3'-AGCATCG^TCGAGTGCCAA |
| I-CreI | Chlamydomonas reinhardtii | 5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG<br>3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC |
| I-ChuI | Chlamydomonas humicola | 5'-GAAGGTTTGGCACCTCG^ATGTCGGCTCATC<br>3'-CTTCCAAACCGTG^GAGCTACAGCCGAGTAG |
| I-CpaI | Chlamydomonas pallidostigma tica | 5'-CGATCCTAAGGTAGCGAA^ATTCA<br>3'-GCTAGGATTCCATC^GCTTTAAGT |
| I-CpaII | Chlamydomonas pallidostigma tica | 5'-CCCGGCTAACTC^TGTGCCAG<br>3'-GGGCCGAT^TGAGACACGGTC |
| I-CeuI | Chlamydomonas | 5'-CGTAACTATAACGGTCCTAA^GGTAGCGAA |

| | eugametos | 3'-GCATTGATATTGCCAG^GATTCCATCGCTT |
|---|---|---|
| I-DmoI | Desulfuro-coccus mobilis | 5'-ATGCCTTGCCGGGTAA^GTTCCGGCGCGCAT<br>3'-TACGGAACGGCC^CATTCAAGGCCGCGCGTA |
| I-SceI | S.cerevisiae | 5'-AGTTACGCTAGGGATAA^CAGGGTAATATAG<br>3'-TCAATGCGATCCC^TATTGTCCCATTATATC<br>   5'-TAGGGATAA^CAGGGTAAT<br>   3'-ATCCC^TATTGTCCCATTA ("Core"-Sequence) |
| I-SceII | S.cerevisiae | 5'-TTTTGATTCTTTGGTCACCC^TGAAGTATA<br>3'-AAAACTAAGAAACCAG^TGGGACTTCATAT |
| I-SceIII | S.cerevisiae | 5'-ATTGGAGGTTTTGGTAAC^TATTTATTACC<br>3'-TAACCTCCAAAACC^ATTGATAAATAATGG |
| I-SceIV | S.cerevisiae | 5'-TCTTTTCTCTTGATTA^GCCCTAATCTACG<br>3'-AGAAAAGAGAAC^TAATCGGGATTAGATGC |
| I-SceV | S.cerevisiae | 5'-AATAATTTTCT^TCTTAGTAATGCC<br>3'-TTATTAAAAGAAGAATCATTA^CGG |
| I-SceVI | S.cerevisiae | 5'-GTTATTTAATG^TTTTAGTAGTTGG<br>3'-CAATAAATTACAAAATCATCA^ACC |
| I-SceVII | S.cerevisiae | 5'-TGTCACATTGAGGTGCACTAGTTATTAC |
| PI-SceI | S.cerevisiae | 5'-ATCTATGTCGGGTGC^GGAGAAAGAGGTAAT<br>3'-TAGATACAGCC^CACGCCTCTTTCTCCATTA |
| F-SceI | S.cerevisiae | 5'-GATGCTGTAGGC^ATAGGCTTGGTT<br>3'-CTACGACA^TCCGTATCCGAACCAA |
| F-SceII | S.cerevisiae | 5'-CTTTCCGCAACA^GTAAAATT<br>3'-GAAAGGCG^TTGTCATTTTAA |
| I-HmuI | Bacillus subtilis bacteriophage SPO1 | 5'-AGTAATGAGCCTAACGCTCAGCAA<br>3'-TCATTACTCGGATTGC^GAGTCGTT |
| I-HmuII | Bacillus subtilis bacteriophage SP82 | 5'-<br>AGTAATGAGCCTAACGCTCAACAANNNNNNNNNNNNNNNNNNNNNNNNN<br>NNNNNNNNNNNNNNNNNNN |
| I-LlaI | Lactococcus lactis | 5'-CACATCCATAAC^CATATCATTTTT<br>3'-GTGTAGGTATTGGTATAGTAA^AAA |
| I-MsoI | Monomastix species | 5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG<br>3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC |
| I-NanI | Naegleria | 5'-AAGTCTGGTGCCA^GCACCCGC |

| | andersoni | 3'-TTCAGACC^ACGGTCGTGGGCG |
|---|---|---|
| I-NitI | Naegleria italica | 5'-AAGTCTGGTGCCA^GCACCCGC |
| | | 3'-TTCAGACC^ACGGTCGTGGGCG |
| I-NjaI | Naegleria jamiesoni | 5'-AAGTCTGGTGCCA^GCACCCGC |
| | | 3'-TTCAGACC^ACGGTCGTGGGCG |
| I-PakI | Pseuden-doclonium akinetum | 5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG |
| | | 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC |
| I-PorI | Pyrobaculum organotrophum | 5'-GCGAGCCCGTAAGGGT^GTGTACGGG |
| | | 3'-CGCTCGGGCATT^CCCACACATGCCC |
| I-PpoI | Physarum polycephalum | 5'-TAACTATGACTCTCTTAA^GGTAGCCAAAT |
| | | 3'-ATTGATACTGAGAG^AATTCCATCGGTTTA |
| I-ScaI | Saccharomyces capensis | 5'-TGTCACATTGAGGTGCACT^AGTTATTAC |
| | | 3'-ACAGTGTAACTCCAC^GTGATCAATAATG |
| I-Ssp6803I | Synechocystis species | 5'-GTCGGGCT^CATAACCCGAA |
| | | 3'-CAGCCCGAGTA^TTGGGCTT |
| PI-PfuI | Pyrococcus furiosus Vc1 | 5'-GAAGATGGGAGGAGGG^ACCGGACTCAACTT |
| | | 3'-CTTCTACCCTCC^TCCCTGGCCTGAGTTGAA |
| PI-PfuII | Pyrococcus furiosus Vc1 | 5'-ACGAATCCATGTGGAGA^AGAGCCTCTATA |
| | | 3'-TGCTTAGGTACAC^CTCTTCTCGGAGATAT |
| PI-PkoI | Pyrococcus kodakaraensis KOD1 | 5'-GATTTTAGAT^CCCTGTACC |
| | | 3'-CTAAAA^TCTAGGGACATGG |
| PI-PkoII | Pyrococcus kodakaraensis KOD1 | 5'-CAGTACTACG^GTTAC |
| | | 3'-GTCATG^ATGCCAATG |
| PI-PspI | Pyrococcus, sp. | 5'-AAAATCCTGGCAAACAGCTATTAT^GGGTAT |
| | | 3'-TTTTAGGACCGTTTGTCGAT^AATACCCATA |
| PI-TfuI | Thermococcus fumicolans ST557 | 5'-TAGATTTTAGGT^CGCTATATCCTTCC |
| | | 3'-ATCTAAAA^TCCAGCGATATAGGAAGG |
| PI-TfuII | Thermococcus fumicolans ST557 | 5'-TAYGCNGAYACN^GACGGYTTYT |
| | | 3'-ATRCGNCT^RTGNCTGCCRAARA |
| PI-ThyI | Thermococcus hydro-thermalis | 5'-TAYGCNGAYACN^GACGGYTTYT |
| | | 3'-ATRCGNCT^RTGNCTGCCRAARA |
| PI-TliI | Thermococcus litoralis | 5'-TAYGCNGAYACNGACGG^YTTYT |
| | | 3'-ATRCGNCTRTGNC^TGCCRAARA |
| PI-TliII | Thermococcus | 5'-AAATTGCTTGCAAACAGCTATTACGGCTAT |

|  | litoralis |  |
|---|---|---|
| I-TevI | Bacteriophage T4 | 5'-AGTGGTATCAAC^GCTCAGTAGATG<br>3'-TCACCATAGT^TGCGAGTCATCTAC |
| I-TevII | Bacteriophage T4 | 5'-GCTTATGAGTATGAAGTGAACACGT^TATTC<br>3'-CGAATACTCATACTTCACTTGTG^CAATAAG |
| F-TevI | Bacteriophage T4 | 5'-GAAACACAAGA^AATGTTTAGTAAANNNNNNNNNNNNNN<br>3'-CTTTGTGTTCTTTACAAATCATTTNNNNNNNNNNNNNN^ |
| F-TevII | Bacteriophage T4 | 5'-TTTAATCCTCGCTTC^AGATATGGCAACTG<br>3'-AAATTAGGAGCGA^AGTCTATACCGTTGAC |

[0099]    Also encompassed are minor deviations (degenerations) of the recognition sequence which still enable recognition and cleavage by the DSBI enzyme in question. Such deviations - also in connection with different framework conditions such as, for example, calcium or magnesium concentration – have been described (Argast GM et al. (1998) J Mol Biol 280: 345-353). Also encompassed are core sequences of these recognition sequences. It is known that the inner portions of the recognition sequences suffice for an induced double-strand break and that the outer ones are not absolutely relevant, but can codetermine the cleavage efficacy. Thus, for example, an 18bp core sequence can be defined for I-SceI.

[00100]    Recombination construct and DSBI enzyme can be combined to give one of the recombination systems or methods according to the invention in various ways with which the skilled worker is familiar. Thus, the recombination constructs and the DSBI enzyme can be combined in an organism, a cell, a cell compartment or a tissue for example as follows:

[00101]    1.) Organisms which have the recombination cassette inserted into the chromosomal DNA are generated in the customary manner. For example, such plants can be generated preferably by agrobacteria-mediated transformation. The primary transformants which contain the recombination cassette are employed for the transformation with an expression cassette which ensures expression of the DSBI enzyme or grown in a suitable manner until homozygous, when they act as the host organism (for example host plant) for the transformation with an expression cassette which ensures expression of the

DSBI enzyme. Starting from these host plants, it is possible, for example, to initiate, establish and use for transformation purposes in-vitro cultures such as, for example, callus cultures or embryogenic cultures. Transformation with the expression cassette for the DSBI enzyme can be in each case stable or transient.

[00102]   2.) Organisms known as master organisms, which bear and express the corresponding gene for the DSBI enzyme (or an expression cassette which ensures the expression of the DSBI enzyme) are generated in the customary manner. For example, such master plants can be generated preferably by agrobacterium-mediated transformation. The primary transformants which express the DSBI enzyme are employed for the transformation with the recombination construct  or grown in a suitable manner until homozygous, which is when they act as master organism or host organism (for example master plant) into which the recombination constructs are introduced. Starting from these master plants, it is possible, for example, to initiate, establish and use for transformation purposes in-vitro cultures such as, for example, callus cultures or embryogenic cultures.

[00103]   3.) The gene encoding the DSBI enzyme (or an expression cassette which ensures the expression of the DSBI enzyme) can be integrated into a vector which already bears the recombination cassette and thus introduced into plant cells simultaneously with the target gene. It is preferred to insert the gene encoding the DSBI enzyme between the homology sequences and thus to delete it from the chromosomal DNA after it has fulfilled its function. Very especially preferably, expression of the DSBI enzyme is inducible in such a case (for example under the control of one of the inducible promotors described herein), in a development-dependent fashion using a development-dependent promotor, or else DSBI enzymes are employed whose activity is inducible in order to avoid cleaving the recombination construct immediately after the transformation and prior to its insertion into the genome.

[00104]   4.) Relying on the co-transformation technique, the expression cassette which ensures the expression of the DSBI enzyme can be transformed into the cells simultaneously with the recombination construct, but on a separate vector. Co-

transformation can be in each case stable or transient. In such a case, expression of the DSBI enzyme is preferably inducible (for example under the control of one of the inducible promotors described herein), in a development-dependent fashion using a development-dependent promotor, or else DSBI enzymes are employed whose activity is inducible in order to avoid cleaving the recombination construct immediately after the transformation and prior to its insertion into the genome.

[00105]     5.) Organisms, for example plants or else animals, expressing the DSBI enzyme may also act as parent individuals. In the progeny from the hybridization between organisms expressing the DSBI enzyme on the one hand and organisms bearing the recombination construct on the other hand, the desired double-strand breaks and recombination between the homology sequences are observed, with the possible deletion of the sequences located between the homology sequences.

[00106]     6.) Expression of the DSBI enzyme is also conceivable in a transient transformation approach in which the possibilities 2 to 4 can be exploited.

[00107]     7.) The DSBI enzyme can also be introduced into cells comprising or bearing the transgenic recombination construct directly, for example via microinjection, particle bombardment (biolistic method), polyethylene glycol transfection or liposome-mediated transfection. This embodiment is advantageous since no DSBI-enzyme-encoding sequences can remain in the genome. Such a method has been described for example by Segal DJ et al. (1995) Proc Natl Acad Sci USA 92:806-810.

[00108]     8.) The DSBI enzyme may also be generated by introducing the DSBI-enzyme-encoding, in-vitro-generated mRNA into cells (for example via microinjection, particle bombardment (biolistic method) or liposome-mediated transfection). This embodiment is advantageous since no DSBI-enzyme-encoding sequences can remain in the genome.

[00109]     9.) The DSBI enzyme can be introduced into plant cells as a fusion protein with the VirE2 or VirF protein of an agrobacterium. Such methods have been described for example for Cre recombinase (Vergunst AC et al. (2000) Science. 290: 979-982). If the expression cassette for the fusion protein is located outside the border sequences, it is not

va-53897

inserted into the plant genome. This embodiment is advantageous since no DSBI-enzyme-encoding sequences can remain in the genome.

[00110]   The recombination system or method according to the invention can be realized in intact organisms or else in parts, cells or propagation material derived therefrom, especially preferably in intact plants or else in any plant tissue or plant in-vitro cultures including callus. An in-vitro application using, for example, wheat germ extract or reticulocyte extract can also be envisaged.

[00111]   As described above, the DSBI enzyme can be generated using an expression cassette which comprises the DNA encoding a DSBI enzyme and is introduced into a eukaryotic cell or organism. In this context, the expression cassette for the DSBI enzyme preferably comprises a nucleic acid sequence encoding a DSBI enzyme as shown in SEQ ID NO: 2, 4, 6, 8 or 10 or a functional equivalent of same which is capable of generating DNA double-strand breaks in double-stranded DNA using the essentially identical recognition sequence. Essentially identical recognition sequences refer to those recognition sequences which, while deviating from the recognition sequence identified as optimal for the enzyme in question, still permit cleavage by this enzyme. Very especially preferably, the expression cassettes for the DSBI enzyme comprise a nucleic acid sequence as shown in SEQ ID NO: 1, 3, 5, 7 or 9, or combinations thereof.

[00112]   Expression cassette – for example when referring to the expression cassette for the DSBI enzyme – means those constructions in which the DNA to be expressed is linked operably to at least one genetic control element which enables or regulates its expression (i.e. transcription and/or translation). Here, expression may be for example stable or transient, constitutive or inducible. For introducing it, the skilled worker may resort to various direct methods (for example transfection, particle bombardment, microinjection) or indirect methods (for example infection with agrobacteria, infection with viruses), all of which are detailed further below.

[00113]   Operable linkage is generally understood as meaning an arrangement in which a genetic control sequence is capable of exerting its function with regard to a nucleic acid

va-53897

sequence, for example while encoding a DSBI enzyme. Function, in this context, may mean for example control of the expression, i.e. transcription and/or translation, of the nucleic acid sequence, for example one encoding a DSBI enzyme. Control, in this context, encompasses for example initiating, increasing, governing or suppressing the expression, i.e. transcription and, if appropriate, translation. Controlling, in turn, may be, for example, tissue- and/or time-specific. It may also be inducible, for example by certain chemicals, stress, pathogens and the like.

[00114] Operable linkage is understood as meaning for example the sequential arrangement of a promotor, of the nucleic acid sequence to be expressed - for example one encoding a DSBI enzyme – and, if appropriate, further regulatory elements such as, for example, a terminator, in such a way that each of the regulatory elements can fulfill its function when the nucleic acid sequence - for example one encoding a DSBI enzyme – is expressed.

[00115] This does not necessarily require a direct linkage in the chemical sense. Genetic control sequences such as, for example, enhancer sequences are also capable of exerting their function on the target sequence from positions located at a distance or indeed other DNA molecules. Preferred arrangements are those in which the nucleic acid sequence to be expressed - for example one encoding a DSBI enzyme – is positioned after a sequence acting as promotor so that the two sequences are linked covalently to one another. The distance between the promotor sequence and the nucleic acid sequence – for example one encoding a DSBI enzyme – is preferably less than 200 base pairs, especially preferably less than 100 base pairs, very especially preferably less than 50 base pairs.

[00116] The skilled worker is familiar with a variety of ways in order to obtain such an expression cassette. For example, it is preferably prepared by directly fusing a nucleic acid sequence which acts as promotor with a nucleotide sequence to be expressed - for example one encoding a DSBI enzyme. Operable linkage can be achieved by means of customary recombination and cloning techniques as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor

va-53897

26

Laboratory, Cold Spring Harbor, NY (1989), in T.J. Silhavy, M.L. Berman and L.W.
Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring
Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology,
Greene Publishing Assoc. and Wiley Interscience (1987).

[00117]    However, an expression cassette may also be constructed in such a way that the
nucleic acid sequence to be expressed (for example one encoding a DSBI enzyme) is
brought under the control of an endogenous genetic control element, for example a
promotor, for example by means of homologous recombination or else by random insertion.
Such constructs are likewise understood as being expression cassettes for the purposes of
the invention.

[00118]    The skilled worker furthermore knows that nucleic acid molecules may also be
expressed using artificial transcription factors of the zinc finger protein type (Beerli RR et
al. (2000) Proc Natl Acad Sci USA 97(4):1495-500). These factors can be adapted to suit
any sequence region and enable expression independently of certain promotor sequences.

[00119]    The term "genetic control sequences" is to be understood in the broad sense and
refers to all those sequences which affect the coming into existence, or the function, of the
expression cassette according to the invention. For example, genetic control sequences
ensure transcription and, if appropriate, translation in prokaryotic or eukaryotic organisms.
Preferably, the expression cassettes according to the invention encompass 5'-upstream of
the respective nucleic acid sequence to be expressed a promotor and 3'-downstream a
terminator sequence as additional genetic control sequence, and, if appropriate, further
customary regulatory elements, in each case in operable linkage with the nucleic acid
sequence to be expressed.

[00120]    Genetic control sequences are described, for example, in "Goeddel; Gene
Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA
(1990)" or "Gruber and Crosby, in: Methods in Plant Molecular Biology and
Biotechnology, CRC Press, Boca Raton, Florida, eds.:Glick and Thompson, Chapter 7, 89-
108" and the references cited therein.

va-53897

[00121] Examples of such control sequences are sequences to which inductors or repressors bind and thus regulate the expression of the nucleic acid. The natural regulation of the sequences before the actual structural genes may still be present in addition to these novel control sequences or instead of these sequences and, if appropriate, may have been genetically modified in such a way that the natural regulation has been switched off and gene expression increased. However, the expression cassette may also be simpler in construction, that is to say no additional regulatory signals are inserted before the abovementioned genes, and the natural promotor together with its regulation is not removed. Instead, the natural control sequence is mutated in such a way that regulation no longer takes place and gene expression is increased. These modified promotors may also be placed on their own before the natural genes for increasing the activity.

[00122] A variety of control sequences are suitable, depending on the host organism or starting organism described in greater detail herein, which, owing to the introduction of the expression cassettes or vectors, becomes a genetically modified, or transgenic, organism.

[00123] Advantageous control sequences for the expression cassettes or vectors according to the invention are present for example in promotors such as cos, tac, trp, tet, phoA, tat, lpp, lac, lacIq, T7, T5, T3, gal, trc, ara, SP6, ?-PR or in the ?-PL promotor, which are advantageously used in Gram-negative bacteria.

[00124] Further advantageous control sequences are present for example in the Gram-positive promotors amy and SPO2, in the yeast or fungal promotors ADC1, MFa , AC, P-60, CYC1, GAPDH, TEF, rp28, ADH or in the plant promotors CaMV/35S (Franck et al. (1980) Cell 21:285-294), PRP1 (Martini N et al. (1993) Mol Gen Genet. 236(2-3):179-186), SSU, OCS, LEB4, USP, STLS1, B33, NOS; FBPaseP (WO 98/18940) or in the ubiquitin or phaseolin promotors.

[00125] Vectors which are suitable for expression in vertebrates, preferably in mammals, are vectors like the TK promotor, the RSV 3' LTR promotor, the CMV promotor or the SV40 early or late promotor. The skilled worker is familiar with other promotors. Inducible promotors suitable for use in vertebrates, preferably in mammals, encompass for example

va-53897

the Tet promotor/repressor, which is inducible or repressible by tetracycline or derivatives, the dexamethasone-inducible MMTV-LTR promotor, the Drosophila minimal heat shock promotor, which is inducible by ecdysone or the analog ponasterone A (for example within the pVgRXR expression system; Invitrogen, Inc.).

[00126] A preferred promotor is, in principle, any promotor which is capable of controlling the expression of genes, in particular foreign genes, in plants. Preferred promotors are those which enable constitutive expression in plants (Benfey et al. (1989) EMBO J. 8:2195-2202). A promotor which is preferably used is, in particular, a plant promotor or a promotor derived from a plant virus. Especially preferred is the promotor of the cauliflower mosaic virus 35S transcript (Franck et al. (1980) Cell 21:285-294; Odell et al.(1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. 1986, Plant Mol. Biol. 6, 221-228) or the 19S CaMV promotor (US 5,352,605 and WO 84/02913). It is known that this promotor comprises a variety of recognition sequences for transcriptional effectors which, in their totality, bring about permanent and constitutive expression of the gene introduced (Benfey et al. (1989) EMBO J. 8:2195-2202). A further suitable constitutive promotor is the Rubisco small subunit (SSU) promotor (US 4,962,028). A further example of a suitable promotor is the leguminB promotor (GenBank Acc.-No.: X03677). Further preferred constitutive promotors are, for example, the Agrobacterium nopaline synthase promotor, the TR dual promotor, the agrobacterium OCS (octopine synthase) promotor, the ubiquitin promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), the promotors of the vacuolar ATPase subunits, or the promotor of a wheat proline-rich protein (WO 91/13991).

[00127] The Expression cassettes may also comprise an inducible, preferably chemically inducible, promotor (Aoyama T and Chua NH (1997) Plant J 11:605-612; Caddick MX et al. (1998) Nat. Biotechnol 16:177-180; Review: Gatz, Annu Rev Plant Physiol Plant Mol Biol 1997, 48:89-108), by means of which the expression of the exogenous gene in the plant can be controlled at a specific point in time. Such promotors, such as, for example, the PRP1 promotor (Ward et al., Plant. Mol. Biol. 22 (1993), 361-366), a salicylic-acid-

inducible promotor (WO 95/19443), a benzenesulfonamide-inducible promotor (EP-A-0388186), a tetracycline-inducible promotor (Gatz et al., (1992) Plant J. 2, 397-404), an abscisic acid-inducible promotor (EP-A 335528), a salicylic acid-inducible promotor (WO 95/19443) or an ethanol- (Salter MG et al. (1998) *Plant J.* 16:127-132) or cyclohexanone-inducible (WO 93/21334) promotor may likewise be used.

[00128] In an especially preferred embodiment, the DSBI-enzyme-encoding nucleic acid, in particular, is expressed under the control of an inducible promotor. This leads to a controlled, governable expression and deletion - for example in plants -, and any problems caused by a constitutive expression of a DSBI enzyme are avoided.

[00129] Other preferred promotors are promotors induced by biotic or abiotic stress, such as, for example, the pathogen-inducible promotor of the PRP1 gene (Ward et al., Plant Mol Biol 1993, 22: 361-366), the tomato heat-inducible hsp80 promotor (US 5,187,267), the potato chill-inducible alpha-amylase promotor (WO 96/12814) or the wound-induced pinII promotor (EP375091).

[00130] Other preferred promotors are promotors with specificity for the anthers, ovaries, pollen, the meristem, flowers, leaves, stems, roots and seeds.

[00131] A development-regulated promotor is, inter alia, described by Baerson et al. (Baerson SR, Lamppa GK (1993) Plant Mol Biol 22(2):255-67).

[00132] Especially preferred promotors are those which ensure expression in tissues or plant parts in which the biosynthesis of starch and/or oils or their precursors takes place or in which the products are advantageously accumulated. The biosynthesis site of starch are the chloroplasts of the leaves or the amyloplasts of the storage organs such as seeds, fruits or tubers. Within these organs, it is predominantly the cells of the endosperm or the cotyledons of the embryo in which synthesis takes place. Preferred promotors are thus in addition to the abovementioned constitutive promotors in particular seed-specific promotors such as, for example, the phaseolin promotor (US 5,504,200, Bustos MM et al., Plant Cell. 1989;1(9):839-53), the promotor of the 2S albumin gene (Joseffson LG et al.(1987) J Biol Chem 262: 12196-12201), the legumin promotor (Shirsat A et al. (1989) Mol Gen Genet.

215(2):326-331), the USP (unknown seed protein) promotor (Bäumlein H et al. (1991) Molecular & General Genetics 225(3):459-67), the napin gene promotor (US 5,608,152; Stalberg K, et al. (1996) L. Planta 199: 515-519), the sucrose binding protein promotor (WO 00/26388) or the legumin B4 promotor (LeB4; Bäumlein H et al. (1991) Mol Gen Genet 225:121-128; Baeumlein et al. (1992) Plant Journal 2(2):233-239; Fiedler U et al. (1995) Biotechnology (NY) 13(10):1090-1093), the Ins Arabidopsis oleosin promotor (WO9845461), the Brassica Bce4 promotor (WO 91/13980). Further suitable seed-specific promotors are those of the genes encoding the "high-molecular-weight glutenin" (HMWG), gliadin, branching enzyme, ADP-glucose pyrophosphatase (AGPase) or starch synthase. Furthermore preferred promotors are those which enable seed-specific expression in monocots such as maize, barley, wheat, rye, rice and the like. Promotors which may advantageously be employed are the promotor of the lpt2 or lpt1 gene (WO 95/15389, WO 95/23230) or the promotors described in WO 99/16890 (promotors of the hordein gene, the glutelin gene, the oryzin gene, the prolamine gene, the gliadin gene, the glutelin gene, the zein gene, the kasirin gene or the secalin gene).

[00133] Promotors which are preferred as genetic control elements are, furthermore, pollen-specific promotors such as, for example, the promotor of the B. campestris bgp1 gene (GenBank Acc.-No: X68210; Xu H et al. (1993) Mol Gen Genet 239(1-2):58-65; WO 94/13809), of the Oryza sativa ory s1 gene (GenBank Acc.-No.: AJ012760; Xu H et al. (1995) Gene 164 (2):255-259), of the pollen-specific maize gene ZM13 (Hamilton DA et al. (1998) Plant Mol Biol 38(4):663-669; US 5,086,169), of the B.napus gene Bp10 (GenBank Acc.-No.: X64257; Albani D (1992) Plant J 2(3):331-342; US 6,013,859), and functional combinations of such promoters.

[00134] Other preferred promotors are the Lcg1 promotor for cell-specific expression in the male gametes (WO 99/05281; XU H et al. (1999) Proc. Natl. Acad. Sci. USA Vol. 96:2554-2558) and the promotor of the AtDMC1 gene (Klimyuk VI et al.(1997) Plant J. 11(1):1-14).

va-53897

[00135]    Further suitable promotors are, for example, specific promotors for tubers, storage roots or roots such as, for example, the class I patatin promotor (B33), the potato cathepsin D inhibitor promotor, the starch synthase (GBSS1) promotor or the sporamin promotor, and fruit-specific promotors such as, for example, the tomato fruit-specific promotor (EP-A 409625).

[00136]    Promotors which are furthermore suitable are those which ensure leaf-specific expression. Promotors which may be mentioned are the potato cytosolic FBPase promotor (WO 98/18940), the Rubisco (ribulose-1,5-bisphosphate carboxylase) SSU (small subunit) promoter, the potato ST-LSI promotor (Stockhaus et al. (1989) EMBO J 8(9):2445-2451) or functional combinations of such promoters. Other preferred promotors are those which govern expression in seeds and plant embryos.

[00137]    Further suitable promotors are, for example, fruit-maturation-specific promotors such as, for example, the tomato fruit-maturation-specific promotor (WO 94/21794), flower-specific promotors such as, for example, the phytoene synthase promotor (WO 92/16635) or the promotor of the P-rr gene (WO 98/22593) or another node-specific promotor as described in EP-A 249676 may be used advantageously.

[00138]    In principle, all natural promotors together with their regulatory sequences, such as those mentioned above, may be used for the method according to the invention. In addition, synthetic promotors may also be used advantageously.

[00139]    Genetic control sequences also encompass further promotors, promotor elements or minimal promotors capable of modifying the expression-specific characteristics. Thus, for example, the tissue-specific expression may take place in addition as a function of certain stress factors, owing to genetic control sequences. Such elements are, for example, described for water stress, abscisic acid (Lam E and Chua NH (1991) J Biol Chem 266(26):17131 -17135) and heat stress (Schoffl F et al. (1989) Molecular & General Genetics 217(2-3):246-53).

[00140]    Furthermore, other promotors which enable expression in further plant tissues or other organisms, such as, for example, E.coli bacteria, may be linked operably with the

va-53897

nucleic acid sequence to be expressed. Plant promotors which are suitable are, in principle, all of the above-described promotors.

[00141] Genetic control sequences furthermore also encompass the 5'-untranslated region, introns or the noncoding 3'-region of genes. It has been demonstrated that they may play a significant role in the regulation of gene expression. Thus, it has been demonstrated that 5'-untranslated sequences are capable of enhancing the transient expression of heterologous genes. Furthermore, they may promote tissue specificity (Rouster J et al., Plant J. 1998, 15: 435-440.). Conversely, the 5'-untranslated region of the opaque-2 gene suppresses expression. Deletion of the region in question leads to an increased gene activity (Lohmer S et al., Plant Cell 1993, 5:65-73).

[00142] Genetic control sequences may also encompass ribosome binding sequences for initiating translation. This is preferred in particular when the nucleic acid sequence to be expressed does not provide suitable sequences or when they are not compatible with the expression system.

[00143] The expression cassette can advantageously comprise one or more of what are known as enhancer sequences in operable linkage with the promotor, which enable the increased transgenic expression of the nucleic acid sequence. Additional advantageous sequences, such as further regulatory elements or terminators, may also be inserted at the 3' end of the nucleic acid sequences to be expressed recombinantly. One or more copies of the nucleic acid sequences to be expressed recombinantly may be present in the gene construct.

[00144] Genetic control sequences are furthermore understood as meaning sequences which encode fusion proteins consisting of a signal peptide sequence.

[00145] Polyadenylation signals which are suitable as genetic control sequences are plant polyadenylation signals, preferably those which correspond essentially to T-DNA polyadenylation signals from *Agrobacterium tumefaciens*, in particular of gene 3 of the T-DNA (octopine synthase) of the Ti plasmids pTiACHS (Gielen et al., EMBO J. 3 (1984), 835 et sec.) or functional equivalents thereof. Examples of particularly suitable terminator

sequences are the OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator.

[00146] As mentioned above, the recombination constructs according to the invention may encompass further nucleic acid sequences. Such nucleic acid sequences may preferably constitute expression cassettes. The following may be mentioned by way of example of the DNA sequences to be expressed in the expression constructs, but not by way of limitation:

[00147]   i)   Positive selection markers:

[00148] As a rule, selection markers are required for selecting cells which have successfully undergone homologous recombination or transformation. The selectable marker which has been introduced together with the expression construct confers resistance to a biocide (for example a herbicide such as phosphinothricin, glyphosate or bromoxynil), a metabolism inhibitor such as 2-deoxyglucose-6-phosphate (WO 98/45456) or an antibiotic such as, for example, tetracyclines, ampicillin, kanamycin, G 418, neomycin, bleomycin or hygromycin to the cells which have successfully undergone recombination or transformation. The selection marker permits the selection of the transformed cells from untransformed cells (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Especially preferred selection markers are those which confer resistance to herbicides. Examples of selection markers which may be mentioned are:

[00149]   -   DNA sequences which encode phosphinothricin acetyltransferases (PAT), which acetylates the free amino group of the glutamine synthase inhibitor phosphinothricin (PPT) and thus brings about detoxification of the PPT (de Block et al. 1987, EMBO J. 6, 2513-2518) (also referred to as Bialophos®resistance gene (bar)),

[00150]   -   5-enolpyruvylshikimate-3-phosphate synthase genes (EPSP synthase genes), which confer resistance to Glyphosate® (N-(phosphonomethyl)glycine),

[00151]   -   the gox gene, which encodes the Glyphosate®-degrading enzyme (Glyphosate oxidoreductase),

[00152]   -   the deh gene (encoding a dehalogenase which inactivates Dalapon®),

[00153]   -   acetolactate synthases which inactivate sulfonylurea and imidazolinone,

va-53897

34

[00154] - bxn genes which encode Bromoxynil®-degrading nitrilase enzymes,

[00155] the kanamycin, or G418, resistance gene (NPTII). The NPTII gene encodes a neomycin phosphotransferase which reduces the inhibitory effect of kanamycin, neomycin, G418 and paromomycin owing to a phosphorylation reaction,

[00156] - the DOG$^R$1 gene. The DOG$^R$1 gene has been isolated from the yeast Saccharomyces cerevisiae (EP 0 807 836). It encodes a 2-deoxyglucose-6-phosphate phosphatase which confers resistance to 2-DOG (Randez-Gil et al. 1995, Yeast 11, 1233-1240).

[00157] ii) Negative selection markers enable for example the selection of organisms with successfully deleted sequences which encompass the marker gene (Koprek T et al. (1999) The Plant Journal 19(6):719-726). TK thymidine kinase (TK) and diphtheria toxin A fragment (DT-A), codA gene encoding a cytosine deaminase (Gleve AP et al. (1999) Plant Mol Biol. 40(2):223-35; Pereat RI et al. (1993) Plant Mol. Biol 23(4): 793-799; Stougaard J; (1993) Plant J 3:755-761), the cytochrome P450 gene (Koprek et al. (1999) Plant J. 16:719-726), genes encoding a haloalkane dehalogenase (Naested H (1999) Plant J. 18:571-576), the iaaH gene (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) or the tms2 gene (Fedoroff NV & Smith DL 1993, Plant J 3: 273- 289).

[00158] iii) Report genes which encode readily quantifiable proteins and which, via intrinsic color or enzyme activity, ensure the assessment of the transformation efficacy or of the location or timing of expression. Very especially preferred here are genes encoding reporter proteins (see also Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) such as

[00159] - "green fluorescence protein" (GFP) (Chui WL et al., Curr Biol 1996, 6:325-330; Leffel SM et al., Biotechniques. 23(5):912-8, 1997; Sheen et al.(1995) Plant Journal 8(5):777-784; Haseloff et al.(1997) Proc Natl Acad Sci USA 94(6):2122-2127; Reichel et al.(1996) Proc Natl Acad Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell Rep 16:267-271; WO 97/41228).

[00160] - Chloramphenicol transferase,

va-53897

[00161]    -    luciferase (Millar et al., Plant Mol Biol Rep 1992 10:324-414; Ow et al. (1986) Science, 234:856-859); permits the detection of bioluminescence,

[00162]    -    ß–galactosidase, encodes an enzyme for which a variety of chromogenic substrates are available,

[00163]    -    ß-glucuronidase (GUS) (Jefferson et al., EMBO J. 1987, 6, 3901-3907) or the uidA gene, which encodes an enzyme for a variety of chromogenic substrates,

[00164]    -    R locus gene product: protein which regulates the production of anthocyanin pigments (red coloration) in plant tissue and thus makes possible the direct analysis of the promotor activity without the addition of additional adjuvants or chromogenic substrates (Dellaporta et al., In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282, 1988),

[00165]    -    ß-lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), enzyme for a variety of chromogenic substrates (for example PADAC, a chromogenic cephalosporin),

[00166]    -    xylE gene product (Zukowsky et al. (1983) Proc Natl Acad Sci USA 80:1101-1105), catechol dioxygenase capable of converting chromogenic catechols,

[00167]    -    alpha-amylase (Ikuta et al. (1990) Bio/technol. 8:241-242),

[00168]    -    tyrosinase (Katz et al.(1983) J Gene Microbiol 129:2703-2714), enzyme which oxidizes tyrosine to give DOPA and dopaquinone which subsequently form melanine, which is readily detectable,

[00169]    -    aequorin (Prasher et al.(1985) Biochem Biophys Res Commun 126(3):1259-1268), can be used in the calcium-sensitive bioluminescence detection.

[00170]    -    The recombination construct according to the invention and any vectors derived from it may comprise further functional elements. The term "further functional elements" is to be understood in the broad sense. It preferably refers to all those elements which affect the generation, multiplication, function, use or value of the recombination system according to the invention, recombination construct according to the invention or

va-53897

36

cells or organisms comprising them. The following may be mentioned by way of example, but not by limitation, of the further functional elements.

[00171]    iv) Replication origins which ensure replication of the expression cassettes or vectors according to the invention in, for example, E. coli. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2$^{nd}$ ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

[00172]    v)  Multiple cloning regions (MCS) enable and facilitate the insertion of one or more nucleic acid sequences.

[00173]    vi) Sequences which make possible homologous recombination or insertion into the genome of a host organism.

[00174]    vii)Elements, for example border sequences, which make possible the agrobacterium-mediated transfer in plant cells for the transfer and integration into the plant genome, such as, for example, the right or left border of the T-DNA or the vir region.

[00175]    All of the abovementioned expression cassettes or further functional elements may be located, as mentioned, between the homology sequences A and B. However, they may also be located outside them. This is advantageous in particular in the case of border sequences.

[00176]    A recombination cassette or expression construct according to the invention for a DSBI enzyme may advantageously be introduced into cells using vectors into which these constructs or cassettes are inserted. Examples of vectors include, for example, plasmids, cosmids, phages, viruses, retroviruses or agrobacteria.

[00177]    Vectors for expression in E.coli are preferably pQE70, pQE60 and pQE-9 (QIAGEN, Inc.); pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Biotech, Inc.), or combinations thereof.

[00178]    Preferred vectors for eukaryotic expression encompass pWLNE0, pSV2CAT, pOG44, pXT1 and pSG (Stratagene Inc.); pSVK3, pBPV, pMSG and pSVL (Pharmacia

va-53897

Biotech, Inc.). Inducible vectors which may be mentioned are pTet-Thia, Potter-Splice, pcDNA4/TO, pcDNA4/TO /LacZ, pcDNA6/TR, pcDNA4/TO/Myc-His /LacZ, pcDNA4/TO/Myc-His A, pcDNA4/TO/Myc-His B, pcDNA4/TO/Myc-His C, pVgRXR (Invitrogen, Inc.), the pMAM series (Clontech, Inc.; GenBank Accession No.: U02443), or combinations thereof. These already provide the inducible regulatory control element for example for a chemical inducible expression of a DSBI enzyme. The nucleic acid sequence encoding a DSBI enzyme may be inserted directly into these vectors.

[00179]    Vectors for the expression in yeast encompass by way of example pYES2, pYD1, pTEFl/Zeo, pYES2/GS, pPICZ,pGAPZ, pGAPZalph, pPIC9, pPIC3.5, PHIL-D2, PHIL-Sl, pPIC3SK, pPIC9K, PA0815 (Invitrogen, Inc.) and combinations thereof.

[00180]    In an advantageous embodiment, the expression cassette is introduced by means of plasmid vectors. Preferred vectors are those which enable the stable integration of the expression cassette into the host genome.

[00181]    Another subject matter of the invention relates to eukaryotic transgenic organisms comprising the recombination system according to the invention, and to cells, cell cultures, tissues, parts or propagation material – such as, for example, in the case of plant organisms leaves, roots, seeds, fruit, pollen and the like – derived from such organisms.

[00182]    Eukaryotic organism, starting organism or host organism refers to higher and lower, single- and multi-celled eukaryotic organisms. Also encompassed are eukaryotic microorganisms such as, for example, yeasts, algae or fungi. Preferred yeasts are Candida, Saccharomyces, Hansenula or Pichia, with Saccharomyces cerevisiae or Pichia pastoris (ATCC Accession No. 201178) being particularly preferred. Preferred fungi are Aspergillus, Trichoderma, Ashbya, Neurospora, Fusarium, Beauveria or other fungi described in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) on page 15, Table 6. The filamentous Hemiascomycete Ashbya gossypii is particularly preferred.

[00183]    Host or starting organisms which are preferred in accordance with the invention are, furthermore, animal organisms and cells or tissue derived from them. Animal

va-53897

organisms encompass preferably vertebrates and invertebrates. Especially preferred vertebrates are mammals such as in dogs, cats, sheep, goats, chickens, mice, rats, bovines or horses. Preferred animal cells encompass CHO, COS and HEK293 cells. Preferred invertebrates encompass insect cells such as Drosophila S2 and Spodoptera Sf9 or Sf21 cells.

[00184]  Host organisms or starting organisms which are preferred as transgenic organisms are especially plants. Included within the scope of the invention are all genera and species of higher and lower plants of the plant kingdom. Included are furthermore the mature plants, seed, shoots and seedlings, and parts, propagation material (for example seeds and fruit) and cultures, for example cell cultures, derived therefrom. Mature plants are to be understood as meaning plants at any developmental stage beyond the seedling. Seedling is to be understood as meaning a young, immature plant in an early developmental stage.

[00185]  The recombination system according to the invention may preferably be used for the following plant families: Amaranthaceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cucurbitaceae, Labiatae, Leguminosae-Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Saxifragaceae, Scrophulariaceae, Solanacea, Tetragoniacea and transgene combinations thereof.

[00186]  Annual, perennial, monocotyledonour and dicotyledonous plants are preferred host organisms for the generation of transgenic plants. The use of the recombination system, or method according to the invention is furthermore advantageous in all ornamental plants, useful or ornamental trees, flowers, cut flowers, shrubs or turf. Plants which may be mentioned by way of example but not by limitation are angiosperms, bryophytes such as, for example, Hepaticae (hepaticas) and Musci (mosses); pteridophytes such as ferns, horsetail and clubmosses; gymnosperms such as conifers, cycads, ginkgo and Gnetaeae; algae such as Chlorophyceae, Phaeophpyceae, Rhodophyceae, Myxophyceae, Xanthophyceae, Bacillariophyceae (diatoms) and Euglenophyceae.

[00187]    Plants for the purposes of the invention comprise by way of example and not by way of limitation the families of the Rosaceae such as rose, Ericaceae such as rhododendrons and azaleas, Euphorbiaceae such as poinsettias and croton, Caryophyllaceae such as pinks, Solanaceae such as petunias, Gesneriaceae such as African violet, Balsaminaceae such as touch-me-not, Orchidaceae such as orchids, Iridaceae such as gladioli, iris, freesia and crocus, Compositae such as marigold, Geraniaceae such as geraniums, Liliaceae such as drachaena, Moraceae such as ficus, Araceae such as philodendron and many others.

[00188]    Flowering plants which may be mentioned by way of example but not by limitation are the families of the Leguminosae such as pea, alfalfa and soya; Gramineae such as rice, maize, wheat; Solanaceae such as tobacco and many others; the family of the Umbelliferae, particularly the genus Daucus (very particularly the species carota (carrot)) and Apium (very particularly the species graveolens dulce (celery)) and many others; the family of the Solanacea, particularly the genus Lycopersicon, very particularly the species esculentum (tomato) and the genus Solanum, very particularly the species tuberosum (potato) and melongena (aubergine) and many others; and the genus Capsicum, very particularly the species annum (peppers) and many others; the family of the Leguminosae, particularly the genus Glycine, very particularly the species max (soybean) and many others; and the family of the Cruciferae, particularly the genus Brassica, very particularly the species napus (oilseed rape), campestris (beet), oleracea cv Tastie (cabbage), oleracea cv Snowball Y (cauliflower) and oleracea cv Emperor (broccoli); and the genus Arabidopsis, very particularly the species thaliana and many others; the family of the Compositae, particularly the genus Lactuca, very particularly the species sativa (lettuce) and many others.

[00189]    The transgenic plants according to the invention are selected in particular among monocotyledonous crop plants, such as, for example, cereals such as wheat, barley, sorghum and millet, rye, triticale, maize, rice or oats, and sugar cane. The transgenic plants according to the invention are furthermore selected in particular from among

va-53897

dicotyledonous crop plants such as, for example, Brassicaceae such oilseed rape, cress, Arabidopsis, cabbages or canola, Leguminosae such as soya, alfalfa, peas, beans or peanut. Solanaceae such as potato, tobacco, tomato, aubergine or peppers, Asteraceae such as sunflower, Tagetes, lettuce or Calendula. Cucurbitaceae such as melon, pumpkin/squash or courgette, and linseed, cotton, hemp. Flax, red pepper, carrot, sugar beet and the various tree, nut and wine species.

[00190]    Especially preferred are Arabidopsis thaliana, Nicotiana tabacum and oilseed rape and all genera and species which are used as food or feedstuffs, such as the above-described cereal species, or which are suitable for the production of oils, such as oil crops (such as, for example, oilseed rape),·nut species, soya, sunflower, pumpkin/squash and peanut.

[00191]    Plant organisms are furthermore, for the purposes of the invention, other organisms which are capable of photosynthetic activity, such as, for example, algae or cyanobacteria, and also mosses. Preferred algae are green algae, such as, for example, algae of the genus Haematococcus, Phaedactylum tricornatum, Volvox or Dunaliella.

[00192]    The generation of a transformed organism or a transformed cell requires introducing the DNA in question into the host cell in question. A multiplicity of methods is available for this procedure, which is termed transformation (see also Keown et al. 1990 Methods in Enzymology 185:527-537). For example, the DNA can be introduced directly by microinjection or by bombardment with DNA-coated microparticles. Also, the cell can be permeabilized chemically, for example using polyethylene glycol, so that the DNA can enter the cell by diffusion. The DNA can also be introduced by protoplast fusion with other DNA-containing units such as minicells, cells, lysosomes or liposomes. Another suitable method of introducing DNA is electroporation, where the cells are permeabilized reversibly by an electrical pulse. Preferred general methods which may be mentioned are the calcium-phosphate-mediated transfection, the DEAE-dextran-mediated transfection, the cationic lipid-mediated transfection, electroporation, transduction and infection. Such methods are

va-53897

known to the skilled worker and described, for example, in Davis et al., Basic Methods In Molecular Biology (1986).

[00193] In plants, methods for transforming and regenerating plants from plant tissues or plant cells with which the skilled worker is familiar are exploited for transient or stable transformation. Suitable methods are especially protoplast transformation by means of polyethylene-glycol-induced DNA uptake, biolistic methods such as the gene gun ("particle bombardment" method), electroporation, the incubation of dry embryos in DNA-containing solution, sonication and microinjection, and the transformation of intact cells or tissues by micro- or macroinjection into tissues or embryos, tissue electroporation, incubation of dry embryos in DNA-containing solution or vacuum infiltration of seeds. In the case of injection or electroporation of DNA into plant cells, the plasmid used need not meet any particular requirement. Simple plasmids such as those of the pUC series may be used. If intact plants are to be regenerated from the transformed cells, the presence of an additional selectable marker gene on the plasmid is useful.

[00194] Any plant tissue may act as target material. Likewise, expression may take place in callus, embryogenic tissue or somatic embryos.

[00195] In addition to these "direct" transformation techniques, transformation can also be carried out by bacterial infection by means of Agrobacterium tumefaciens or Agrobacterium rhizogenes. These strains contain a plasmid (Ti or Ri plasmid). Part of this plasmid, termed T-DNA (transferred DNA), is transferred to the plant following agrobacterial infection and integrated into the genome of the plant cell.

[00196] The recombination construct or the expression cassette for the DSBI enzyme is preferably integrated into specific plasmids, either into a shuttle, or intermediate, vector or into a binary vector). If, for example, a Ti or Ri plasmid is to be used for the transformation, at least the right border, but in most cases the right and the left border, of the Ti or Ri plasmid T-DNA is linked with the expression cassette to be introduced as a flanking region. Binary vectors are preferably used. Binary vectors are capable of replication both in E. coli and in Agrobacterium. As a rule, they contain a selection marker gene and a linker or

va-53897

polylinker flanked by the right or left T-DNA flanking sequence. They can be transformed directly into Agrobacterium (Holsters et al., Mol. Gen. Genet. 163 (1978), 181-187). The selection marker gene permits the selection of transformed agrobacteria and is, for example, the nptII gene, which imparts resistance to kanamycin. The agrobacterium, which acts as host organism in this case, should already contain a plasmid with the vir region. The latter is required for transferring the T-DNA to the plant cell. An agrobacterium thus transformed can be used for transforming plant cells.

[00197] The use of Agrobacterium tumefaciens for the transformation of plants using tissue culture explants has been described by Horsch et al. (Horsch RB (1986) Proc Natl Acad Sci USA 83(8):2571-2575), Fraley et al. (Fraley et al. 1983, Proc. Natl. Acad. Sci. USA 80, 4803-4807) and Bevans et al. (Bevans et al. 1983, Nature 304, 184-187). Many strains of Agrobacterium tumefaciens are capable of transferring genetic material - for example the recombination constructs according to the invention -, such as, for example, the strains EHA101[pEHA101], EHA105[pEHA105], LBA4404[pAL4404], C58C1[pMP90] and C58C1[pGV2260]. The strain EHA101[pEHA101] has been described by Hood et al. (Hood EE et al. (1996) J Bacteriol 168(3):1291-1301), the strain EHA105[pEHA105] by Hood et al. (Hood et al. 1993, Transgenic Research 2, 208-218), the strain LBA4404[pAL4404] by Hoekema et al. (Hoekema et al. 1983, Nature 303, 179-181), the strain C58C1[pMP90] by Koncz and Schell (Koncz and Schell 1986, Mol. Gen. Genet. 204, 383-396), and the strain C58C1[pGV2260] by Deblaere et al. (Deblaere et al. 1985, Nucl. Acids Res. 13, 4777-4788).

[00198] The agrobacterial strain employed for the transformation comprises, in addition to its disarmed Ti plasmid, a binary plasmid with the T-DNA to be transferred, which, as a rule, comprises a gene for the selection of the transformed cells and the gene to be transferred. Both genes must be equipped with transcriptional and translational initiation and termination signals. The binary plasmid can be transferred into the agrobacterial strain for example by electroporation or other transformation methods (Mozo & Hooykaas 1991,

va-53897

43

Plant Mol. Biol. 16, 917-918). Coculture of the plant explants with the agrobacterial strain is usually performed for two to three days.

[00199] A variety of vectors could, or can, be used. In principle, one differentiates between those vectors which can be employed for the agrobacterium-mediated transformation or agroinfection, i.e. which comprise the recombination constructs, or the expression cassette, for the expression of the DSBI enzyme within a T-DNA, which indeed permits stable integration of the T-DNA into the plant genome. Moreover, border-sequence-free vectors may be employed, which can be transformed into the plant cells for example by particle bombardment, where they can lead both to transient and to stable expression.

[00200] The use of T-DNA for the transformation of plant cells has been studied and described intensively (EP 120516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B. V., Alblasserdam, Chapter V; Fraley et al., Crit. Rev. Plant. Sci., 4:1-46 and An et al., EMBO J. 4 (1985), 277-287). Various binary vectors are known, some of which are commercially available such as, for example, pBIN19 (Clontech Laboratories, Inc. USA).

[00201] To transfer the DNA to the plant cell, plant explants are cocultured with Agrobacterium tumefaciens or Agrobacterium rhizogenes. Starting from infected plant material (for example leaf, root or stalk sections, but also protoplasts or suspensions of plant cells), intact plants can be regenerated using a suitable medium which may contain, for example, antibiotics or biocides for selecting transformed cells. The plants obtained can then be screened in the presence of the DNA introduced, in this case the recombination construct or the expression cassette for the DSBI enzyme according to the invention. As soon as the DNA has integrated into the host genome, the genotype in question is, as a rule, stable and the insertion in question is also found in the subsequent generations. As a rule, the expression cassette integrated contains a selection marker which confers a resistance to a biocide (for example a herbicide) or an antibiotic such as kanamycin, G 418, bleomycin, hygromycin or phosphinotricin and the like to the transformed plant. The selection marker permits the selection of transformed cells (McCormick et al., Plant Cell Reports 5 (1986),

81-84). The plants obtained can be cultured and hybridized in the customary fashion. Two or more generations should be grown in order to ensure that the genomic integration is stable and hereditary.

[00202]    The abovementioned methods are described, for example, in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), 128 - 143 and in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225). The construct to be expressed is preferably cloned into a vector which is suitable for the transformation of Agrobacterium tumefaciens, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711).

[00203]    Agrobacterium-mediated transformation is suited best to dicotyledonous plant cells, whereas the direct transformation techniques are suitable for any cell type.

[00204]    Transformed cells, i.e. those which comprise the DNA integrated into the DNA of the host cell, can be selected from untransformed cells if a selectable marker is part of the DNA introduced. A marker can be, for example, any gene which is capable of conferring a resistance to antibiotics or herbicides. Transformed cells which express such a marker gene are capable of surviving in the presence of concentrations of a suitable antibiotic or herbicide which kill an untransformed wild type. Various positive and negative selection markers are described hereinabove. Examples are the bar gene, which confers resistance to the herbicide phosphinothricin (Rathore KS et al.,Plant Mol Biol. 1993 Mar;21(5):871-884), the nptII gene, which confers resistance to kanamycin, the hpt gene, which confers resistance to hygromycin, or the EPSP gene, which confers resistance to the herbicide Glyphosate.

[00205]    As soon as a transformed plant cell has been generated, an intact plant can be obtained using methods known to the skilled worker. For example, callus cultures are used as starting material. The formation of shoot and root can be induced in this as yet undifferentiated cell biomass in the known fashion. The shoots obtained can be planted and cultured.

[00206] Also in accordance with the invention are cells, cell cultures, parts – such as, for example, in the case of transgenic plant organisms, roots, leaves and the like – derived from the above-described transgenic organisms, and transgenic propagation material (such as seeds or fruits).

[00207] Genetically modified plants according to the invention which can be consumed by humans or animals can also be used as food or feedstuffs, for example directly or following processing known per se. Here, the deletion of, for example, resistances to antibiotics and/or herbicides, as are frequently introduced when generating the transgenic plants, makes sense for reasons of customer acceptance, but also product safety.

[00208] A further subject matter of the invention relates to the use of the above-described transgenic organisms according to the invention and the cells, cell cultures, parts – such as, for example, in the case of transgenic plant organisms, roots, leaves and the like – derived from them, and transgenic propagation material such as seeds or fruits, for the production of food or feedstuffs, pharmaceuticals or fine chemicals. Here again, the deletion of, for example, resistances to antibiotics and/or herbicides is advantageous for reasons of customer acceptance, but also product safety.

[00209] Fine chemicals is understood as meaning enzymes, vitamins, amino acids, sugars, fatty acids, natural and synthetic flavors, aromas and colorants widely usable. Especially preferred is the production of tocopherols and tocotrienols, and of carotenoids. Culturing the transformed host organisms, and isolation from the host organisms or from the culture medium, is performed by methods known to the skilled worker. The production of pharmaceuticals such as, for example, antibodies or vaccines, is described by Hood EE, Jilka JM. (1999) Curr Opin Biotechnol. 10(4):382-386; Ma JK and Vine ND (1999) Curr Top Microbiol Immunol.236:275-92).

[00210] The recombination system or method according to the invention furthermore offers various advantageous uses which can not be achieved with the deletion methods described in the prior art. Various use examples are described herein by way of example, but not by limitation:

va-53897

46

**[00211]** Simple deletion of a nucleic acid sequence from the chromosomal DNA of an organism:

**[00212]** Using any homology sequences A and B, nucleic acid sequences located between them can be deleted. The sequence which is the result of the recombination of the homology sequences A and B remains in the genome. The method is suitable for example for removing, from the chromosomal DNA, selection markers after a transgenic organism, for example a transgenic plant, has been generated. The method is shown schematically in Figs. 2 and 3, Fig. 2 showing the variant with one recognition sequence for the site-directed induction of DNA double-strand breaks and Fig. 3 showing the variant with two recognition sequences for the site-directed induction of DNA double-strand breaks.

**[00213]** Complete deletion of recombinantly introduced heterologous nucleic acid sequences from the chromosomal DNA of an organism:

**[00214]** Using homology sequences A and B, which are homologous to certain sequences of the organism, the expression construct can be introduced into the organism by homologous recombination. Using the recombination system or method according to the invention, the nucleic acid sequences located between the homology sequences would be deleted. The induced homologous recombination between homology sequences A and B restores the original sequence. All of the construct is removed from the chromosomal DNA. The method is suitable for example for removing selection markers from the chromosomal DNA after a transgenic plant has been generated. Furthermore, the system or method according to the invention is suitable for expressing certain proteins transiently in order to achieve an advantageous effect and then to switch them off using an induced DSBI enzyme expression or activity by irreversibly removing the gene in question from the genome. The method is shown schematically in Fig. 4, the variant with two recognition sequences for the site-directed induction of DNA double-strand breaks being shown. The system can also be realized using one recognition sequence; however, two cleavage sites are advantageous in the case of larger insertions between the homology sequences A and B since this allows the

deletion efficacy and homologous recombination efficacy to be increased further (further recognition sequences may be located within the sequence region to be deleted).

[00215] Induced gene activation by the site-directed deletion of nucleic acid sequences:

[00216] Using homology sequences A and B, whose homologous recombination restores for example a complete open reading frame of a protein, or a functional promotor, the inducible expression of target proteins can be realized as a function of the presence of the DSBI enzyme. Using the recombination system or method according to the invention, the nucleic acid sequences located between the homology sequences would be deleted. The method is shown schematically in Figs. 5 and 6, Fig. 6 showing a specific embodiment of the general method shown in Fig. 5 in which the recombination construct is inserted into an endogenous gene at an earlier point in time by means of homologous recombination, thus enabling this gene to be inducibly activated as a function of the presence of the DSBI enzyme. Fig. 7a illustrates the system of gene activation with reference to a specific use example where the ß-glucuronidase (GUS) gene is reconstituted using the system or method according to the invention, enabling a color reaction to take place (see description to Fig. 7a and Examples).

[00217] Readily selectable system for the deletion of a nucleic acid sequence from the chromosomal DNA of an organism:

[00218] - In a preferred embodiment, the recombination construct comprises a positive and a negative selection marker (and, if appropriate, further nucleic acid sequences to be deleted) in such a way that both markers are deleted when the double-strand breaks are induced. Such a system is shown in Figs. 8 and 9 (A). Moreover, the expression cassette for the DSBI enzyme may also be present between the homology sequences (Fig. 10 (B)), expression preferably being effected under the control of an inducible promotor (Pi) (for example: Aoyama T and Chua NH (1997) Plant J 11:605-612; Caddick MX et al. (1998) Nat. Biotechnol 16:177-180). As already described, further nucleic acid sequences may be present (Fig. 9 (C)).

[00219] - Expression of the DSBI enzyme leads in both cases to the elimination of the DNA sequences located between the two recognition sequences and the recombination of the homologous sequences. Since the cells simultaneously lose a negative selection marker, the cells with a successful deletion can be identified by means of selection (Gleave AP et al.(1999) Plant Mol Biol. 40:223-235).

[00220] - In the case of plant cells, for example, the resulting cells can be used for regenerating and propagating the corresponding intact plants, which now no longer contain any marker genes.

[00221] Genetic manipulation of the host genome:

[00222] The recombination system or method according to the invention can be used for in-situ modifications of the host genome. Thus, for example, a homology sequence may already exist endogenously in the genome. After insertion of the second homology sequence, which is linked with a DSBI enzyme recognition sequence, any regulatory or coding sequences located between the homology sequences A and B are eliminated from the genome.

[00223] At the same time, it is conceivable that the recombination construct encompasses regulatory or coding sequences which are eliminated from the organism once deletion has taken place. Thus, it is possible for example to regulate transiently an endogenous gene in a site-directed fashion.

[00224] - In a further preferred embodiment, the efficacy of the recombination system is increased by combination with systems which promote homologous recombination. Such systems are described and encompass, for example, the expression of proteins such as RecA or the treatment with PARP inhibitors. It has been demonstrated that the intrachromosomal homologous recombination in tobacco plants can be increased by using PARP inhibitors (Puchta H et al. (1995) Plant J. 7:203-210). Using these inhibitors, the homologous recombination rate in the recombination constructs after induction of the sequence-specific DNA double-strand break, and thus the efficacy of the deletion of the transgenesequences, can be increased further. Various PARP inhibitors may be employed for this purpose.

va-53897

49

Preferably encompassed are inhibitors such as 3-aminobenzamide, 8-hydroxy-2-methylquinazolin-4-one (NU1025), 1,11b-dihydro-[2H]benzopyrano[4,3,2-de]isoquinolin-3-one (GPI 6150), 5-aminoisoquinolinone, 3,4-dihydro-5-[4-(1-piperidinyl)butoxy]-1(2H)-isoquinolinone, or the compounds described in WO 00/26192, WO 00/29384, WO 00/32579, WO 00/64878, WO 00/68206, WO 00/67734, WO 01/23386 and WO 01/23390.

[00225]    In addition, it was possible to increase the frequency of various homologous recombination reactions in plants by expressing the E. coli RecA gene (Reiss B et al. (1996) Proc Natl Acad Sci USA 93(7):3094-3098). Also, the presence of the protein shifts the ratio between homologous and illegitimate DSB repair in favor of homologous repair (Reiss B et al. (2000) Proc Natl Acad Sci USA 97(7):3358-3363). Reference may also be made to the methods described in WO 97/08331 for increasing the homologous recombination in plants. A further increase in the efficacy of the recombination system might be achieved by the simultaneous expression of the RecA gene or other genes which increase the homologous recombination efficacy (Shalev G et al. (1999) Proc Natl Acad Sci USA 96(13):7398-402). The above-stated systems for promoting homologous recombination can also be advantageously employed in cases where the recombination construct is to be introduced in a site-directed fashion into the genome of a eukaryotic organism by means of homologous recombination.

[00226]    Sequences:

1.    SEQ ID NO:1

Nucleic acid sequence for the I-SceI homing endonuclease.

2.    SEQ ID NO:2

Protein sequence for the I-SceI homing endonuclease.

3.    SEQ ID NO:3

Nucleic acid sequence for fusion protein of I-ChuI homing endonuclease and N-terminal nuclear localization sequence.

4.    SEQ ID NO:4

va-53897

Protein sequence for fusion protein of I-ChuI homing endonuclease and N-terminal nuclear localization sequence.

5.  SEQ ID NO:5

Nucleic acid sequence for fusion protein of I-CreI homing endonuclease and N-terminal nuclear localization sequence.

6.  SEQ ID NO:6

Protein sequence for fusion protein of I-CreI homing endonuclease and N-terminal nuclear localization sequence.

7.  SEQ ID NO:7

Nucleic acid sequence for fusion protein of I-CpaI homing endonuclease and N-terminal nuclear localization sequence.

8.  SEQ ID NO:8

Protein sequence for fusion protein of I-CpaI homing endonuclease and N-terminal nuclear localization sequence.

9.  SEQ ID NO:9

Nucleic acid sequence for fusion protein of I-CpaII homing endonuclease and N-terminal nuclear localization sequence.

10.  SEQ ID NO:10

Protein sequence for fusion protein of I-CpaII homing endonuclease and N-terminal nuclear localization sequence.

11.  SEQ ID NO: 11:     oligonucleotide primer OPN1

5'-CGG CTC GAG CTA CGG GGA CGA TTT CTT TTT TTC AC-3'

12.  SEQ ID NO: 12:     oligonucleotide primer OPN2

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC AT GAA TAC AAA ATA TAA TAA AGA GTT CTT ACT C-3'

13.  SEQ ID NO: 13: oligonucleotide primer OPN3

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG GAC ATT AAT CCT CAA TGG ATT ACA GG- 3'

14.     SEQ ID NO: 14: oligonucleotide primer OPN4

5'-CGG CTC GAG TTA CTC GCC AGT TTC TTC AAA ACG-3'

15.     SEQ ID NO: 15: oligonucleotide primer OPN5

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG ACC GAT TCT AAA TCT AGA AAC AAC-3'

16.     SEQ ID NO: 16:      oligonucleotide primer OPN6

5'-CGG CTC GAG CTA AAG GTG GCC TTT ATT GCC ATC AG-3'

17.     SEQ ID NO: 17:      oligonucleotide primer OPN7

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG TCA TTA ACA CAA CAA CAA AAA GAC-3'

18.     SEQ ID NO: 18:      oligonucleotide primer OPN8

5'-CGG CTC GAG CTA AAG GTG GCC TTT ATT GCC ATC AG-3'

19.     SEQ ID NO: 19: oligonucleotide primer OPN9

5'-CGG CTC TAG AGC GGC CGC CTA GGG ATA ACA GGG TAA TAG AAT CCC ACA AAA ATC TGA GCT TAA CAG 3'

20.     SEQ ID NO: 20:      oligonucleotide primer OPN10

5'-CGG CTC TAG ACT ATT ACC CTG TTA TCC CTA GGC CCG ATC TAG TAA CAT AGA TGA CAC CGC GCG CG 3'

21.     SEQ ID NO: 21:      oligonucleotide primer OPN11

5'- CGG AAG CTT CGT CAC CAA TCC CAA TTC GAT CTA C - 3'

22.     SEQ ID NO: 22:      oligonucleotide primer OPN12

5'- CGG AAG CTT CCA CTT GCA AAG TCC CGC TAG TGC C - 3'

23.     SEQ ID NO: 23:      oligonucleotide primer OPN13

va-53897

52

5'- CGG AAG CTT CGT CAC CAA TCC CAA TTC GAT CTA C - 3'

24.    SEQ ID NO: 24:    oligonucleotide primer OPN14

5'- CGG AAG CTT CCA CTT GCA AAG TCC CGC TAG TGC C - 3'

25.    SEQ ID NO: 25:    oligonucleotide primer OPN15

5'- CTA GTA CAA AAC GTC GTG AGA CAT TTT AAT CTG AAG GTT TGG CAC CTC GAT GTC GGC TCA TC-3'

26.    SEQ ID NO: 26:    oligonucleotide primer OPN16

5'-CTA GGA TGA GCC GTC ATC GAG GTG CCA AAC CTT CAG ATT AAA ATG TCT CAC GAC GTT TTG TA-3'

27.    SEQ ID NO: 27:    oligonucleotide primer OPN17

5'-CTA GTC CGA AAA CGC CGT GAG ACA TAT TGG TTA CGA TCC TAA GGT AGC GAA ATT CAC CCG GTA ACT CTG TGC CAG-3'

28.    SEQ ID NO: 28:    oligonucleotide primer OPN18

5'-CTA GCT GGC ACA GAG TTA CCG GGT GAA TTT CGC TAC CTT AGG ATC GTA ACC AAT ATG TCT CAC GGC GTT TTC GGA-3'

29.    SEQ ID NO: 29:    nuclear localization sequence NLS1

N-Pro-lys-Thr-Lys-Arg-Lys-Val-C

30.    SEQ ID NO: 30:    nuclear localization sequence NLS2

N-Pro-Lys-Lys-Lys-Arg-Lys-Val-C (SEQ ID NO: 30)

[00227]    Figures

[00228]    The following abbreviations apply to the figures in general:

H1:    Homology sequence A

H2:    Homology sequence B

H1/2:    Sequence as the result of homologous recombination of H1 and H2

S1:    First recognition sequence for the site-directed induction of DNA double-strand breaks

S2:    Second recognition sequence for the site-directed induction of DNA double-strand breaks.

va-53897

E:      DSBI enzyme

P:      Promotor or other genetic control element

N:      Further nucleic acid sequence

NS:     Negative selection marker

PS:     Positive selection marker

T1:     Front part, for example of a gene or open reading frame

T2:     Back part, for example of a gene or open reading frame

STOP: Interruption of a gene or open reading frame by, for example, stop codons or reading-frame shifts.

[00229]   Fig. 1:      Diagram of the principle of the invention.  Sequences in the genome can be eliminated efficiently when they are flanked by the homology sequences H1 and H2 and when a cleavage site (S1) for a DSBI enzyme is located between the homology sequences. Owing to the action of the DSBI enzyme (E) on this recombination cassette (H1-S1-H2), double-strand breaks are formed at the cleavage site S1 and the sequences located between H1 and H2 are eliminated.

[00230]   Fig. 2:      Preferred embodiment. Sequences – in the present case for example an expression cassette consisting of a promotor (P) and a further nucleic acid sequence (N) to be expressed (for example a selection marker) – can be eliminated efficiently from the chromosomal DNA when the are flanked by the homology sequences H1 and H2 and when a cleavage site (S1) for a DSBI enzyme is located between the homology sequences. Owing to the action of the DSBI enzyme (E) on this recombination cassette (H1-S1-P-N-H2), double-strand breaks are formed at the cleavage site S1 and the sequences located between H1 and H2 are eliminated. The cleavage site S1 may also be located behind or within the expression cassette.

[00231]   Fig. 3:      Preferred embodiment. Sequences – in the present case for example an expression cassette consisting of a promotor (P) and a further nucleic acid sequence (N) to be expressed (for example a selection marker) – can be eliminated particularly efficiently from the chromosomal DNA when they are flanked by the homology sequences H1 and H2

va-53897

54

and when in each case one cleavage site (S1 and S2) for a DSBI enzyme is located before and after the nucleic acid sequence to be deleted. Owing to the action of the DSBI enzyme (E) on this recombination cassette (H1-S1-P-N-S2-H2), double-strand breaks are formed at the cleavage sites S1 and S2 and the sequences located between H1 and H2 are eliminated.

[00232]    Fig. 4:    Preferred embodiment. Sequences – in the present case for example an expression cassette consisting of a promotor (P) and a further nucleic acid sequence (N) to be expressed (for example a selection marker) can be eliminated virtually without trace from the chromosomal DNA when the recombination construct comprising them has previously been inserted into the host genome, for example by homologous recombination. In doing so, the gene consisting of the sequence segments T1, H1/2 and T2 is interrupted. The recombination construct is flanked by two parts of the interrupted gene (T1-H1 or H2-T2), the middle part (H1 or H2) having been duplicated in order to enable homologous recombination to take place. The action of the DSBI enzyme (E) on the cleavage sites (S1 and S2) induces double-strand breaks and induces the homologous recombination between the homology sequences H1 and H2, whereby firstly the sequences located between H1 and H2 are deleted and secondly the original gene is restored.

[00233]    Fig. 5:    Preferred embodiment. Nucleic acid sequences (in the present case a gene with the sequence T1-H1/2-T2 under the control of a promotor P) can be expressed inducibly by reconstituting the intact gene only by applying the recombination system. The gene, consisting of the sequence segments T1, H1/2 and T2 is inactivated, for example by the insertion of stop codons or other interruptions of the reading frame within the recombination construct. The recombination construct is flanked by two parts of the interrupted gene (T1-H1 or H2-T2), the middle part (H1 or H2) having been duplicated in order to enable homologous recombination to take place. The action of the DSBI enzyme (E) on the cleavage sites (S1 and S2) induces double-strand breaks and induces the homologous recombination between the homology sequences H1 and H2, whereby firstly the sequences located between H1 and H2 are deleted and secondly the intact gene is restored.

va-53897

[00234] Fig. 6: Preferred embodiment. The figure shows a method which is identical to that described in Fig. 5, only that in the present case an endogenous gene is to be activated in a site-directed manner by introducing the recombination construct for example by means of homologous recombination.

[00235] Fig. 7a: Exemplary embodiment. The figure illustrates a specific embodiment of the method described in Fig. 6. A recombination construct is introduced via agrobacterium-mediated transfection. Flanked by the right and left border sequence (RB and LB, respectively), the construct comprises the interrupted reading frame of the GUS gene (ß-glucuronidase) under the control of the 35S promotor (P) and the nopaline synthase (nos) terminator. The middle region of the GUS gene (U) was duplicated and constitutes the homology sequences A and B. Located between these sequences is the codA gene as negative selection marker under the control of the Cauliflower Mosaic Virus (CaMV) 35S promotor and the nopaline synthase (nos) terminator, flanked by two recognition sequences of the DSBI enzyme (S1 and S2). The recombination construct furthermore additionally comprises the BAR gene under the control of the 35S promotor (P) and the 35S terminator, as positive selection marker.

[00236] Fig. 7a illustrates the appearance of double-strand breaks and the homologous recombination between the homologous U sequences, brought about by the action of the DSBI enzyme, whereby firstly the sequences located between the homologous U sequences are deleted and secondly the GUS gene is restored. The length of the Acc65I fragment is thus reduced from 7.3 kb to 3.7 kb.

[00237] Fig. 7b: Shows the same system as described under Fig. 7a. Fig. 7a illustrates the appearance of double-strand breaks as the result of the action of the DSBI enzyme. In contrast to Fig. 7a, no homologous recombination takes place in the present case, but illegitimate recombination by non-homologous end-joining. While the region between S1 and S2 is deleted owing to the two cleavage sites, the GUS gene is not restored. The length of the Acc65I fragment is thus reduced from 7.3 kb to 4.4 kb.

va-53897

[00238] Fig. 7c: The figure is another representation of the two end products of the processes described under Fig.7a and Fig.7b. A: Result of the homologous recombination; Acc65I fragment has a length of 3.7 kb; the size of the fragment amplified with the aid of the primers OPN13 and OPN14 (shown by the arrows) is 0.7 kb. B: Result of the illegitimate recombination (non-homologous end-joining); Acc65I fragment has a length of 4.4 kb; the size of the fragment amplified with the aid of the primers OPN13 and OPN14 (shown by the arrows) is 1.4 kb.

[00239] Fig. 8: Preferred embodiment. The recombination cassettes advantageously encompass both a positive and a negative selection marker (PS and NS, respectively), in each case under the control of a promotor. The positive selection marker is useful for facilitating and detecting the introduction of the construct into the genome. The negative selection marker is useful for detecting the deletion of the construct from the genome. Both markers are eliminated efficiently from the chromosomal DNA when they are flanked by the homology sequences H1 and H2 and when in each case one cleavage site (S1 and S2, respectively) for a DSBI enzyme is located before and/or after the nucleic acid sequence to be deleted. Owing to the effect of the DSBI enzyme (E) on this recombination cassette, double-strand breaks appear at the cleavage sites S1 and/or S2 and the sequences located between H1 and H2 are then eliminated.

[00240] The effect of one of the abovementioned DSBI enzymes brings about site-directed double-strand breaks and induces the homologous recombination between the homologous U sequences, whereby firstly the sequences located between the homologous U sequences are deleted and secondly the GUS gene is restored.

[00241] Fig.9: Readily selectable systems for deleting a nucleic acid sequence from the chromosomal DNA of an organism. The constructs comprise a positive selection marker (PS) and negative selection marker (NS), in each case under the control of a promotor (P).

[00242] (B) additionally comprises an expression cassette for the DSBI enzyme, expression preferably taking place under the control of an inducible promotor (Pi). (C) Further nucleic acid sequences may be present.

va-53897

[00243]   Expression of the DSBI enzyme leads in all cases to the elimination of the DNA sequences located between the two recognition sequences and to the recombination of the homologous sequences. Since the cells simultaneously lose a negative selection marker, the cells where successful deletion has taken place can be identified by means of selection (Gleave AP et al.(1999) Plant Mol Biol. 40:223-235).

[00244]   Fig. 10:  The figure illustrates the two constructs (SI construct (A) and SD construct (B)) which were used for proving that homologous recombination by double-strand breaks can be induced with different restriction enzymes. The constructs are introduced via agrobacterium-mediated transfection. The constructs, which are flanked by the right and left border sequence (RB and LB, respectively) contain the interrupted reading frame of the GUS gene (□-glucuronidase) under the control of the 35S promotor (P) and the nopaline synthase (nos) terminator. The middle region of the GUS gene (U) was duplicated and constitutes the homology sequences A and B. Located between these sequences are, in the case of the SI construct (A), the recognition sequences of the DSBI enzymes I-SceI, I-CpaI, I-CpaII and I-CreI, and, in the case of the SD construct (B), the recognition sequence of the I-ChuI enzyme. The recombination constructs furthermore additionally contain the BAR gene under the control of a promotor (P) as positive selection marker.

[00245]   Fig. 11:  Representative histochemical analysis of tobacco calli obtained after the induction of double-strand breaks. A blue coloration (here shown as dark coloration) indicates the expression of the ß-glucuronidase gene, and thus the elimination of the selection marker by homologous recombination. Blue (dark colorations) can be seen in the case of the calli in the wells A2, A5, A6, B2, C1, C6 and D2.

[00246]   Fig. 12:  PCR analysis for detecting homologous recombination. PCR of DNA from tobacco calli using the primers OPN13 and OPN14.  The PCR product (size 0.7 kb) which indicates homologous recombination can be seen in lanes 1, 2 and 3. The corresponding calli were blue following histochemical staining, and the corresponding PCR bands were sequenced in order to demonstrate that the open reading frame (ORF) of ß-

va-53897

glucuronidase was indeed restored by homologous recombination. Lanes 4 and 5: PCR products (1.4 kb) of calli which did not turn blue upon staining, where the transgene was eliminated by non-homologous end-joining.

[00247]   Fig. 13: Southern blots which indicate the complete elimination of the transgene sequence in question. The lanes of blots A to D comprise in each case:

| Lane | Line | Description |
|---|---|---|
| 1 | GU.C.USB 1 | Original line |
| 2 | GU.C.USB 1-61 | Non-homologous end-joining |
| 3 | GU.C.USB 1-83 | Homologous recombination |
| 4 | GU.C.USB 3 | Original line |
| 5 | GU.C.USB 3-1 | Non-homologous end-joining |
| 6 | GU.C.USB 3-3 | Homologous recombination |
| 7 | GU.C.USB 7 | Original line |
| 8 | GU.C.USB 7-14 | Non-homologous end-joining |
| 9 | GU.C.USB 7-34 | Homologous recombination |

A:   HindIII-digested DNA hybridized with a ß-glucuronidase-specific probe.

B:   HindIII-digested DNA hybridized with a codA-specific probe.

C:   Acc65I-digested DNA hybridized with a ß-glucuronidase-specific probe.

D:   Acc65I-digested DNA hybridized with a codA-specific probe.

The analysis demonstrates that, following the induction of DNA double-strand breaks by means of expression of the restriction enzyme, not only homologous recombination (lanes 3, 6 and 9), but also illegitimate recombination (lanes 2, 5 and 8) may occur, the transgene sequence (codA) located between the restriction cleavage sites always having been eliminated from the plant genome.

[00248]   The following examples illustrate embodiments of the invention, but should not be viewed as limiting the scope of the invention.

[00249]   **Examples**

[00250]   General methods:

va-53897

[00251] The chemical synthesis of oligonucleotides can be affected for example in the known manner using the phosphoamidite method (Voet, Voet, 2nd edition, Wiley Press New York, pages 896-897). The cloning steps carried out for the purposes of the present invention, such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, the transfer of nucleic acids to nitrocellulose and nylon membranes, the linkage of DNA fragments, the transformation of E. coli cells, bacterial cultures, the propagation of phages and the sequence analysis of recombinant DNA are carried out as described by Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. Recombinant DNA molecules were sequenced using an ALF Express laser fluorescence DNA sequencer (Pharmacia, Upsala, Sweden) following the method of Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

[00252] Example 1: Cloning of the homing endonucleases

[00253] The open reading frames (ORFs) of the homing endonucleases I-CreI (Wang J et al. (1997) Nucleic Acids Res 25: 3767-3776), I-ChuI (Cote V et al.(1993) Gene 129:69-76), I-CpaI (Turmel M et al. (1995a) Nucleic Acids Res 23:2519-2525) and I-CpaII (Turmel M et al. (1995b) Mol. Biol. Evol. 12, 533-545) were cloned from the corresponding Chlamydomonas strains.

[00254] To ensure optimal translation of the gene, the ORFs of the endonucleases were linked to the leader sequence of a plant virus (CaMV gene V, as has proven useful in the case of I-SceI; Puchta H (1993) Nucl Acids Res 21:5034-5040). Also, a nuclear localization sequence (NLS2; SEQ ID NO: 30) was placed in front of the ORFs in order to efficiently transport the protein to the intended site of action. The two elements (leader sequence and nuclear localization sequence) were introduced via PCR by means of the oligonucleotide primers used.

[00255] To isolate the open reading frames (ORFs) of the endonucleases from Chlamydomonas, the algal cultures Chlamydomonas reinhardtii/Smith (strain no. 11-32b), Chlamydomonas applanata/Lucksch (strain no.: 11-9) and Chlamydomonas segris/King (strain no.: 9.83) were obtained from the algal culture collection at Göttingen (University of

va-53897

60

Göttingen, experimental phycology and collection of algal cultures, Albrecht-von-Haller Institute for Plant Sciences, Untere Karspüle 2, D-37073 Göttingen). The cultures were grown with the aid of a shake culture in MS medium, and DNA was obtained using the DNeasy Plant Maxi Kit (Qiagen, Hilden).

[00256] The I-CreI ORF (GenBank Acc. No.: X01977) was amplified from a sample of the algal culture 11-32b Chlamydomonas reinhardtii/Smith with the aid of the oligonucleotides OPN1 and OPN2 (SEQ ID NO: 11 and 12).

[00257] OPN1 (SEQ ID NO: 11):

5'-CGG CTC GAG CTA CGG GGA CGA TTT CTT TTT TTC AC- 3'

OPN2 (SEQ ID NO: 12):

5'- CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC AT GAA TAC AAA ATA TAA TAA AGA GTT CTT ACT C 3'

2 ml (corresponding to approximately 100 ng DNA) of the DNA preparation were employed in the PCR reaction. The following were combined in a total volume of 50 ml in accordance with the manufacturer's instructions (Life Technologies):

[00258] --5 ml 10X PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl] 1.5 ml 50 mM $MgCl_2$ 1 ml 10 mM dNTP mix (10 mM of each dATP, dCTP, dGTP and dTTP) 1 ml primer OPN1 (10 mM) 1 ml primer OPN2 (10 mM) 0.4 ml Taq DNA polymerase (5 U/ml) 2 ml DNA preparation 38.1 ml autoclaved distilled water.

[00259] - The reaction mixture is covered with approx. 50 ml of silicone oil and subjected to the following temperature program (Thermocycler: MWG Biotech Primus HT; MWG Biotech, Germany):

1      cycle of 180 sec at 95°C

30      cycles of 60 sec at 92°C, 60 sec at 54°C and 3 min at 72°C

1      cycle of 5 min at 72°C.

[00260] - The PCR fragment was purified via agarose gel electrophoresis using the QIAquickr Gel Extraction Kit (Qiagen, Hilden, Germany) and cloned into the pGEM-T

va-53897

Easy vector (Promega, Madison, USA). Next, a sequence analysis is carried out using the ALF-Express DNA sequencer (Pharmacia, Upsala, Sweden). The sequence is shown in SEQ ID NO: 5.

[00261] - Cloning of the ORF of I-CpaI from the algal culture 9.83 Chlamydomonas segris/King (Genbank Acc. No.: L36830) was carried out analogously to the description given for I-CreI. The oligonucleotides OPN3 and OPN4 were used for the PCR. The sequence is shown in SEQ ID NO: 7.

[00262] OPN3 (SEQ ID NO: 13):

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG GAC ATT AAT CCT CAA TGG ATT ACA GG- 3'

OPN4 (SEQ ID NO: 14):

5'-CGG CTC GAG TTA CTC GCC AGT TTC TTC AAA ACG-3'

[00263] - Cloning the ORF of I-CpaII was also carried out analogously as described for I-CreI (Genbank Acc. No: L39865). A sample of the algal culture 9.83 Chlamydomonas segris/King was used for this purpose. The oligonucleotides OPN5 and OPN6 were used for the PCR. The sequence is shown in SEQ ID NO: 9.

[00264] OPN5 (SEQ ID NO: 15):

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG ACC GAT TCT AAA TCT AGA AAC AAC-3'

[00265] OPN6 (SEQ ID NO: 16):

5'-CGG CTC GAG CTA AAG GTG GCC TTT ATT GCC ATC AG-3'

[00266] - Cloning of the ORF of I-ChuI from the algal culture 11-9 Chlamydomonas applanata/Lucksch (Genbank Acc. No.: L06107) was carried out analogously to the description given for I-CreI. The oligonucleotides OPN7 and OPN8 were used for the PCR. The sequence is shown in SEQ ID NO: 3.

[00267] OPN7 (SEQ ID NO: 17):

va-53897

62

5'-CGG CTC GAG TAC CTA GAA TAC AAA GAA GAG GAA GAA GAA ACC TCT
ACA GAA GAA GCC ATG GGT CCA AAG AAA AAG AGA AAG GTT ATC ATG
TCA TTA ACA CAA CAA CAA AAA GAC-3'

[0001]    OPN8 (SEQ ID NO: 18):

5'-CGG CTC GAG CTA AAG GTG GCC TTT ATT GCC ATC AG-3')

[00268]    The ORF of the individual homing endonucleases (with the nuclear localization signal) was in each case excised from the pGEM-T Easy vector by means of restriction digestion with SalI, purified by gel electrophoresis and in each case cloned into the SalI restriction cleavage site of the binary vector pBinAR (Höfgen and Willmitzer (1990) Plant Science 66:221-230). Expression of the individual enzymes takes place under the control of the 35S promotor and the octopine synthase terminator.

[00269]    The binary I-*SceI* expression vector pCISceI (Puchta H et al. (1996) Proc. Natl. Acad. Sci. USA 93:5055-5060) contains a synthetic I-*SceI* ORF under the control of the CaMV 35S promotor (Puchta H et al. (1993) Nucl Acids Res 21: 5034-5040) between the T-DNA borders.

[00270]    All of the five plasmids were multiplied in E. coli, purified by means of the QIAfilter Plasmid Midi kit (Qiagen, Hilden) and transferred into the agrobacterial strain C58 by means of electroporation.

[00271]    Example 2: Generation of the construct pGU.I.USB

[00272]    The plasmid pGU.US (Tinland B et al. (1994) Proc. Natl. Acad. Sci. USA 91:8000-8004) was used for constructing the recombination substrates. Within the T-DNA region, the plasmid contain two overlapping halves of the ß-glucuronidase (GUS) gene with a 557 bp overlap. A hygromycin gene is integrated in a unique XbaI cleavage site between the GUS sequences.

[00273]    In a first step, the BAR gene together with promotor and terminator sequences was excised from the vector pRC (Puchta H et al. (1996) Proc Natl Acad Sci USA 93:5055-5060) in the form of an isolated HindIII fragment, separated from the vector sequence via agarose gel electrophoresis, excised from the gel, isolated with the aid of the QIAquickr Gel

va-53897

Extraction Kit (Qiagen, Hilden, Germany) and thereafter inserted into the unique HindIII cleavage site of pGU.US. To this end, the vector pGU.US was previously cut with HindIII and dephosphorylated with alkaline phosphatase (calf intestinal alkaline phosphatase (CIP), New England Biolabs, Frankfurt, Germany) in order to prevent recircularization. The resulting vector is termed pGU.US-BAR.

[00274] In the vector pNE3 (Stougaard J (1993) Plant J 3:755-761), the XbaI cleavage site was first removed by a Klenow filling-in reaction. The open reading frame (ORF) of the negative selection marker gene cytosine deaminase (codA) under the control of the Cauliflower Mosaic Virus (CaMV) 35S promotor and of the nopaline synthase (nos) terminator was amplified from the resulting vector pNE3-XBA by means of PCR using the oligonucleotide primers ONP9 (SEQ ID NO: 16) and ONP10 (SEQ ID NO: 17). Owing to the oligonucleotide primers OPN9 and OPN10 used, in each case one I-*Sce*I cleavage site (emphasized in bold in the sequences stated herein) and a NotI or XbaI cleavage site were added to the two ends of the amplificate.

[00275] OPN9 (SEQ ID NO: 19):

5'-CGG CTC TAG AGC GGC CGC **CTA GGG ATA ACA GGG TAA** TAG AAT CCC ACA AAA ATC TGA GCT TAA CAG 3'

[0001] OPN10 (SEQ ID NO: 20):

5'-CGG CTC TAG ACT **ATT ACC CTG TTA TCC CTA** GGC CCG ATC TAG TAA CAT AGA TGA CAC CGC GCG CG 3'

[00276] - 2 ml (corresponding to approximately 100 ng) of a plasmid preparation of pNE3-XBA were employed for the PCR reaction. The following were combined in a total volume of 50 ml in accordance with the manufacturer's instructions (Life Technologies):

[00277] 5 ml 10X PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl] 1.5 ml 50 mM MgCl$_2$ 1 ml 10 mM dNTP mix (10 mM of each dATP, dCTP, dGTP and dTTP) 1 ml primer OPN1 (10 mM) 1 ml primer OPN2 (10 mM) 0.4 ml Taq DNA polymerase (5 U/ml) 2 ml plasmid preparation of pNE3-XBA 38.1 ml autoclaved distilled water.

[00278] The reaction mixture is covered with approx. 50 ml of silicone oil and subjected to the following temperature program (Thermocycler: MWG Biotech Primus HT; MWG Biotech, Germany):

1    cycle of 180 sec at 95°C

30    cycles of 60 sec at 92°C, 60 sec at 54°C and 3 min at 72°C

1    cycle of 5 min at 72°C.

[00279]    -    The PCR product was digested with XbaI and NotI. The vector pGU-US-BAR was likewise digested with XbaI and NotI (which resulted in the deletion of the hygromycin marker gene), and the vector fragment was purified by agarose gel electrophoresis using the QIAquickr Gel Extraction Kit (Qiagen, Hilden, Germany). Ligation of the digested PCR fragment and vector gave rise to the binary vector pGU.C.USB (see Fig. 7a). The vector contains a marker gene (cytosine deaminase (codA)) on a T-DNA between two I-SceI cleavage sites. The I-SceI cleavage sites are outwardly flanked by homologous sequence regions 557 bp in size of the ß-glucuronidase gene (GUS). The GUS gene acts as homologous restoration marker (Swoboda P et al. (1994) EMBO J 13:481- 489). If the gene is restored by homologous recombination, the expression can be detected histochemically. Elimination of the marker gene gives rise to 5-FC (fluorocytosine)-resistant tobacco cells, which can then regenerate it to give calli (Salomon S and Puchta H (1998) EMBO J 17:6086-6095).

Example 3: Plant transformation with pGU.I.USB

[00280]    -    Nicotiana tabacum L. cv. Petite Havana Line SR1 seedlings were transformed with the agrobacterial strain C58, which contained the binary vector pGU.C.USB.

[00281]    To this end, seeds were placed on moistened filter paper under sterile conditions and the seedlings were harvested after 2 weeks, all as described by Puchta H. (1999) Methods Mol Biol 113: 447-451 (25°C, 16 hours light/8 hours dark rhythm).

[00282]    For the inoculation, the agrobacteral strain containing the binary transformation plasmid was first grown overnight in a shake culture at 28°C in YEB medium. Then, the

va-53897

65

agrobacterial suspension was centrifuged for 10 minutes at 15.000 g and the cells were taken up in 10 mM MgSO₄ so that the final optical density of the suspension had a value of approximately 0.5. In a reaction vessel, the seedlings were then placed into the bacterial solution under sterile conditions and a vacuum of 0.15 at was applied in a sterile desiccator. After 10 minutes, the seedlings were then placed on MS plates supplemented with BAP (6-benzylaminopurine 5 mg/ml) and NAA (1-naphthaleneaceteic acid 0.5 mg/ml) and left for 3 days in a growth cabinet (25°C, 16 hours light/8 hours dark rhythm). The seedlings were then placed on MS medium supplemented additionally with phosphinothricin (100 mg/ml), vancomycin (1 mg/ml) and cefotaxim (0.5 mg/ml) in addition to NAA and BAP. Every 10 days, the seedlings were transferred to freshly made plates. Eventually, the resulting calli formed shoots. As soon as the shoots had attained a certain size (1 to 2 cm), they were excised from the callus material and planted in magenta boxes comprising MS medium supplemented with phosphinothricin, vancomycin and cefotaxin (concentrations as above). After a short time, the shoots developed roots; they were transferred into soil after 2 to 4 weeks. The plants were made to flower in the greenhouse and were then shelled, and the seeds formed were left to ripen in the capsules. The seeds were then placed on MS medium supplemented with 300 mg of phosphinothricin (for the positive selection) or 500 mg of 5-FC (fluorocytosin; for the negative selection) per ml in order to carry out the segregation analyses. By determining the ratio of resistant to sensitive seedlings (3:1 in the case of positive selection and 1:3 in the case of negative selection), it was possible to demonstrate that the recombination constructs were inserted at a locus in the three selected lines.

[00283] Example 5: Induction of gene deletion by introducing the DSBI enzyme I-SceI

[00284] In the experiments, F1 seedlings of the transgenic lines GU.C.USB 1, 3 and 7, each of which comprises a copy of the T-DNA GU.C.USB shown in Fig. 2, were inoculated with an agrobacterial strain which transiently expressed I-SceI and which comprised the plasmid pCISceI (Puchta H et al. (1996) Proc Natl Acad Sci USA 93, 5055-5060) in the above described manner (see also Puchta, 1999b). After 3 days, the seedlings were plated

va-53897

66

on MS medium supplemented with BAP and NAA (concentrations as above) medium to the same medium additionally in the presence of 100 mg of 5-FC and 100 mg of phosphinothricin per ml incubated in order to detect plant cells in which the marker gene to be eliminated (in this case the codA gene) was deleted. After 6 weeks, the calli growing on the medium were divided into two, and one part was used for the regeneration of shoot axes while the other was used for isolating DNA and for the ß-glucuronidase assay. The resulting 5-FC-resistant transgenic calli were analyzed for homologous recombination events by means of histochemical staining. A blue staining indicated restoration of the callus (see Fig. 11).

[00285] The histochemical staining of the calli was carried out as described by Swoboda et al., 1994. To this end, the calli were introduced into staining solution (0.3 mg X-Gluc [Duchefa, Harlem, Nl] per ml of 100 mM sodium phosphate buffer pH 7.0; 0.1% Triton; 0.05% $NaN_3$). A vacuum was applied for 15 minutes to the desiccator, and the calli were subsequently incubated in the solution for 48 hours at 37°C. After the staining solution was poured off, the remaining chlorophyll was removed from the plant material by repeated shaking in 80% ethanol. The blue staining obtained indicated the ß-glucuronidase activity.

[00286] In approximately one quarter of the cases, the marker gene was eliminated successfully by homologous recombination (Fig. 11, Table 2).

[00287] Table 2. Number of 5-FC-resistant tobacco calli following transient DSB induction

| Transgenic line | Seedlings | resis- tant calli | GUS-positive | GUS-positive (% of resistant calli) |
|---|---|---|---|---|
| GU.C.USB 1 | 290 | 56 | 22 | 39 |
| GU.C.USB 3 | 490 | 90 | 24 | 27 |
| GU.C.USB 7 | 370 | 59 | 11 | 19 |

va-53897

[00288] Molecular analyses confirm the facts: since the line GU.C.USB 1 contained a single copy of the transgene, the calli were analyzed directly for recombination results by means of PCR.

[00289] A random fraction of calli was then analyzed at the molecular level by means of PCR. Molecular analysis with the primer pairs

[00290] OPN11 (SEQ ID NO: 21)

5'- CGG AAG CTT CGT CAC CAA TCC CAA TTC GAT CTA C - 3' and

OPN12 (SEQ ID NO: 22)

5'- CGG AAG CTT CCA CTT GCA AAG TCC CGC TAG TGC C - 3'

allowed the isolation of the newly-formed linkage sites from the tobacco genome (Fig. 12; Table 3).

[00291] Table 3. Molecular analysis of recombination events by means of PCR

| Transgenic line | Calli | PCR fragment(s) | | |
|---|---|---|---|---|
| | | 0.7 kb | 1.4 kb | none/other |
| GU.C.USB 1 | 30 | 10 | 12 | 7 |

[00292] - Three 0.7 kb PCR fragments were selected and sequenced. In all three cases, sequencing confirmed the functional sequence of the ß-glucuronidase gene, i.e. the restoration of the gene did indeed take place accurately by homologous recombination.

[00293] When five 1.4 kb PCR bands were sequenced, it was found that these bands were formed after excision of the codA gene by reparation of the two I-SceI cleavage sites (by non-homologous end-joining, NHEJ) without homologous recombination taking place. In most cases, minor deletions at the I-SceI cleavage site resulted.

[00294] Southern blots demonstrated that, as expected, complete elimination of the sequence located between the I-SceI cleavage sites took place in the recombinants with the 0.7 and 1.4 kb bands, respectively. No codA-specific DNA whatsoever was detectable any longer in the genome of the regenerated plants (Fig. 13 B and D, lanes 2 and 3).

va-53897

68

[00295]    The DNA was isolated with the aid of the DNeasy Plant Mini Kit (Quiagen, Hilden). To detect the recombination products, genomic DNA was analyzed by means of PCR using the oligonucleotides OPN13 and OPN14.

[00296]    OPN13 (SEQ ID NO: 23):

5'- CGG AAG CTT CGT CAC CAA TCC CAA TTC GAT CTA C - 3'

OPN14 (SEQ ID NO: 24):

5'- CGG AAG CTT CCA CTT GCA AAG TCC CGC TAG TGC C - 3'

[00297]    5 ml 10X PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl] 1.5 ml 50 mM MgCl$_2$ 1 ml 10 mM dNTP mix (10 mM of each dATP, dCTP, dGTP and dTTP) 1 ml primer OPN1 (10 mM) 1 ml primer OPN2 (10 mM) 0.4 ml Taq DNA polymerase (5 U/ml) 2 ml DNA preparation 38.1 ml autoclaved distilled water.

[00298]    The reaction mixture is covered with approx. 50 ml of silicone oil and subjected to the following temperature program (Thermocycler: MWG Biotech Primus HT; MWG Biotech, Germany):

1    cycle of 180 sec at 95°C

30    cycles of 60 sec at 92°C, 60 sec at 54°C and 3 min at 72°C

1    cycle of 5 min at 72°C.

[00299]    -    The PCR products were sequenced using the "ABI Prism Dye Terminator Cycle Sequencing Reaction Kit" (PE Applied Biosystems, Weiterstadt).

[00300]    -    For the Southern blotting, the DNA was cut with HindIII or Acc65I and subjected to electrophoresis in a 0.8% agarose gel. The DNA in the gel was then transferred to the hybridization membrane 'Hybond N' (Amersham, Little Chalfont, UK) by means of capillary blotting as described in the manufacturer's instructions. For the molecular hybridization, codA- or GUS-specific gene fragments were isolated from the starting plasmids (XbaI/XhoI fragment as PNE3; Stougaard, 1993, and KpnI/SacI fragment from pGUS23, Puchta and Hohn, 1991, isolated using the QIAquick Gel Extraction Kit [Qiagen, Hilden]) and labeled with the aid of a "Random Priming Labeling Kit" (Megaprime DNA

va-53897

69

labeling system RPN1607, Amersham, Little Chalfont, UK) and [□-$^{32}$P]dATP (Amersham, Little Chalfont, UK). The hybridization reactions were carried out at 65°C.

[00301] Since in each case 2 genetically linked transgene copies were integrated in the case of lines GU.C.USB 3 and GU.C.USB 7, a representative number of plants was regenerated from callus in the case of these lines, DNA was obtained and then analyzed per Southern blot (Table 4).

[00302] In the case of Acc65I, the presence of a GUS-specific 3.7 kb band suggests a homologous recombination, while a 4.4 kb band suggests an NHEJ event ("non-homologous end-joining"; NHEJ) (Fig. 7b and c; Fig. 13 C).

[00303] Table 4. Molecular analysis of recombination events by means of Southern blots

| Transgenic line | Calli | Acc65I fragment (kb) | | |
|---|---|---|---|---|
| | | 3.7 | 4.4 | Deletion |
| GU.C.USB 3 | 39 | 6 | 18 | 15 |
| GU.C.USB 7 | 14 | 2 | 5 | 7 |

[00304] - Interestingly, the same type of linkage in the two transgene copies was found in all cases. In other words, either only homologous recombinations or only NHEJ events occurred. In no case did both possibilities exist in parallel, i.e. for example a homologous recombination in the one transgene and an NHEJ event in the other.

[00305] In both lines, PCR analyses were also carried out, and in each case three 0.7 kb PCR fragments were selected and sequenced. In all three cases, sequencing revealed the functional sequence of the ß-glucuronidase gene, i.e. the restoration of the gene did indeed occur by means of homologous recombination.

[00306] When a total of nine 1.4 kb PCR bands of the two lines were sequenced, it was furthermore found that these bands indeed originated after excision of the codA gene by repair of the two I-SceI cleavage sites (by "non-homologous end-joining" NHEJ). Again, minor deletions resulted at the I-SceI cleavage site in most cases.

[00307] Southern blots demonstrated that, as expected, the sequence between the I-SceI cleavage sites were eliminated completely in the recombinant. No codA-specific DNA

va-53897

70

whatsoever was detected any longer in the genome of the regenerated plants (Fig. 13 B and D, lanes 5, 6 and 8, 9).

[00308]   Example 5:

[00309]   Various transgenic tobacco plant lines were generated which, between the halves of the ß-glucuronidase gene (arrangement as described above) also contained cleavage sites for the abovementioned restriction enzymes in addition to a I-SceI cleavage site by means of cloning synthetic oligonucleotides (Fig. 10). Seedlings of this tobacco line were inoculated in each case in direct comparison with agrobacteria capable of expressing either I-SceI or the corresponding enzyme in plant cells. The resulting calli were then stained histochemically after 2 weeks. The results are shown in Table 4.

[00310]   The plasmid pGU.C.US.B was cut with I-SceI so that the codA gene was excised from the plasmid. The digested DNA was separated by means of agarose gel electrophoresis, the larger band was excised and purified by means of the QIAquick Gel Extraction Kit (Qiagen, Hilden) and subsequently ligated and transformed into E. coli. The resulting plasmid was then cut with XbaI.

[00311]   The complementary single-stranded oligonucleotides OPN25 and OPN26 were made double-stranded by briefly heating to 92°C and subsequent cooling and then subsequently ligated with the XbaI-cut plasmid. The resulting SI construct (pSI) contains the cleavage sites for I-SceI, I-CpaI, I-CpaII and I-CreI ((see Fig. 10 (A)).

[00312]   OPN15 (SEQ ID NO: 25):

5'- CTA GTA CAA AAC GTC GTG AGA CAT TTT AAT CTG AAG GTT TGG CAC CTC GAT GTC GGC TCA TC-3'

OPN16 (SEQ ID NO: 26):

5'-CTA GGA TGA GCC GTC ATC GAG GTG CCA AAC CTT CAG ATT AAA ATG TCT CAC GAC GTT TTG TA-3'

[00313]   The complementary single-stranded oligonucleotides OPN27 and OPN28 were made double-stranded by briefly heating to 92°C and subsequent cooling and then

subsequently ligated with the XbaI-cut plasmid. The resulting SD construct (pSD) contains the cleavage sites for I-SceI and I-ChuI (see Fig. 10 (B)).

[00314]    OPN17 (SEQ ID NO: 27):

5'-CTA GTC CGA AAA CGC CGT GAG ACA TAT TGG TTA CGA TCC TAA GGT AGC GAA ATT CAC CCG GTA ACT CTG TGC CAG-3'

OPN18 (SEQ ID NO: 28):

5'-CTA GCT GGC ACA GAG TTA CCG GGT GAA TTT CGC TAC CTT AGG ATC GTA ACC AAT ATG TCT CAC GGC GTT TTC GGA-3'

[00315]    Transgenic tobacco plants with the two constructs were generated as described further above by means of agrobacterium transformation. Lines which only contained transgenic sequences at one locus were used for the further experiments. These lines were determined by the 3:1 segregation into phosphinothricin-resistant and nonresistant plants. The shelled seedlings were then inoculated with agrobacterial strains which comprised one of the four constructs for expressing the restriction endonucleases or, as the vector control, the plasmid BinAR or, as the positive control, a 1:1 mixture of BinAR and CISce-I. The inoculations were carried out as described above (Puchta H (1999) Methods Mol. Biol. 113:447-451), and for selection purposes the seedlings were grown over several weeks on MS medium supplemented with 100 µg kanamycin per ml, which also contained BAP and NAA, vancomycin and cefotaxin (concentrations as above). The resulting calli were then subjected to histochemical ß-glucuronidase staining as described above.

[00316]    All four of the tested restriction enzymes were capable of inducing homologous recombination in the same order of magnitude as I-SceI (which was employed here in a co inoculation with the selection vector pBinAR [AR]) (Table 5). This demonstrates that homologous recombination can be induced efficiently when using any restriction endonucleases.

[00317]    Table 5. Induction of homologous recombination in plants by means of various endonucleases I-CreI, I-CpaI, I-CpaII and I-ChuI. [Sectors/calli] refers to the number of areas stained blue in the resistant calli.

va-53897

72

| Transgenic line | Enzyme | Sectors/calli | Ratio |
|---|---|---|---|
| SI5 | I-SceI/AR | 42/31 | 1.35 |
| | I-CreI | 77/50 | 0.54 |
| | I-CpaII | 51/50 | 1.02 |
| SI2 | I-SceI/AR | 8/9 | 0.89 |
| | I-CreI | 40/18 | 2.22 |
| | I-CpaII | 9/20 | 0.45 |
| SI2 | I-CpaI | 144/106 | 1.36 |
| SD2 | I-ChuI | 166/100 | 1.66 |

[00318]  -  Other embodiments and uses of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. All references cited herein, including all U.S. and foreign patents and patent applications, and all publications or other documentary materials, are specifically and entirely hereby incorporated herein by reference. It is intended that the specification and examples be considered exemplary only, with the true scope and spirit of the invention indicated by the following claims.

va-53897